AD-A086 176 GEORGIA INST OF TECH ATLANTA F/G 5/9
TESTING FOR LEARNING WITH SMALL DATA SETS.(U)
JUN 78 K A YEALY DAAG39-77-C-0041

UNCLASSIFIED NL

1 OF 2
ADA
086176

PHOTOGRAPH THIS SHEET

ADA086176

DTIC ACCESSION NUMBER

II

LEVEL

INVENTORY

Georgia Inst. of Technology
Atlanta, GA.

"Testing for Learning with Small Data Sets"
Master's thesis
June 1978

DOCUMENT IDENTIFICATION

Contract No. DAAG39-77-C-0041

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DISTRIBUTION STATEMENT

ACCESSION FOR	
NTIS GRA&I	☒
DTIC TAB	☐
UNANNOUNCED	☐
JUSTIFICATION	
BY Per DTIC Form 50 on	
DISTRIBUTION / file	
AVAILABILITY CODES	
DIST	AVAIL AND/OR SPECIAL
A	

DISTRIBUTION STAMP

DTIC
S ELECTE D
JUL 1 1980
D

DATE ACCESSIONED

80 6 20 002

DATE RECEIVED IN DTIC

PHOTOGRAPH THIS SHEET AND RETURN TO DTIC-DDA-2

DTIC FORM OCT 79 70A

DOCUMENT PROCESSING SHEET

ADA086176

TESTING FOR LEARNING WITH SMALL DATA SETS

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

by

Kenneth Alan Yealy

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Operations Research

Contract No. DAAG39 77 C [illegible]

Georgia Institute of Technology

June, 1978

TESTING FOR LEARNING WITH SMALL DATA SETS
by
Kenneth Alan Yealy

Errata

Page	Line	Correction
iv	4-8	Add 1 to each page number listed
18	1	the denominator should be $\sqrt{\hat{\sigma}_{\varepsilon}^2/n}$
22	4	derivation, not deviation
29	4	examining, not examin-
31	14	lower limit on the sumation in the numerator should be (N/2) + 1
33	1	$\sum t_i$, not t_i
33	5	least not last
40	6	extend radical to cover (N-1)
41	5	H_o, not H.
41	14	d_{LLSR} = slope of the fitted line (omit what is currently there)
42	6	Insert a period after the first word, capitalize the t on the.
44	3	capital W on with
48	15	change first on to of
49	26	insert the word search after direct
56	9	estimates, not extimates.
56	14	change of to for
58	15	change extimates to estimates
68	11	insert σ_{ε} after for
77		insert N=6 on body of Figure 3-25
78	2	change slop to slope
91	10	add) at end of line
110	11	insert rate at end of line

TESTING FOR LEARNING WITH SMALL DATA SETS

Approved:

Russell G. Heikes

Russell G. Heikes, Chairman

Harrison M. Wadsworth

Harrison M. Wadsworth, Jr.

Leslie G. Callahan Jr.

Leslie G. Callahan, Jr.

Thomas L. Sadosky

Thomas L. Sadosky

Date approved by Chairman 5-30-78

ACKNOWLEDGMENTS

A number of persons have assisted me in conducting this research. To all of them I express my appreciation. In particular, I thank Dr. Russell G. Heikes, the chairman of my thesis committee, for his guidance, encouragement and continual evaluation of the progress of this study. I also thank the members of the thesis committee, Dr. Harrison M. Wadsworth, Jr., Dr. Leslie G. Callahan, Jr., Dr. Thomas L. Sadosky, for their assistance and support throughout the research.

Finally I would like to express special appreciation to my wife, Lynn, for her patience and moral support throughout the duration of this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

The objective of this research was to develop a simple methodology to test for learning using a small sample size and to develop a procedure for measuring the rate of learning at any particular trial. For this research the time between trials was considered insignificant in affecting previously gained knowledge and the error between any observation and its expected value, z_i, is assumed to be NID $(0,\sigma_\varepsilon^2)$.

Assuming learning can be described by a performance curve of the form $z = 1-at^{-b}$ two linear methods and one non-linear method were developed to test for learning by examining the rate of learning over several trials. Since the curve is monotonically increasing a positive slope will be interpreted as learning and a zero slope will correspond to no learning occurring. The linear procedures are based on testing the average rate of learning that occurs over several trials. Several methods for estimating the average rate of learning and the variance of the observations, σ_ε^2, were investigated. The best method for estimating the average rate of learning, based on the minimum variance of the estimate, was the linear least squares regression, LLSR method, and the best estimator of σ_ε^2, which resulted in the most powerful test, was computed using the first differences of the

observations.

In the nonlinear method, estimates for σ_ε and the parameters "a" and "b" are obtained and a test on the degree of nonlinearity of the function is conducted using Beales measure of nonlinearity. If the degree of nonlinearity is small enough then the confidence interval for the slope at any trial can be evaluated by using linear theory approximations.

In a comparison of the two procedures, the linear methods were more powerful tests, however, the nonlinear method was able to provide information on the rate of learning at each trial when the nonlinearity conditions were satisfied and significant learning was detected. The more powerful linear test procedure was the LLSR method, which can detect an average rate of learning over 15 trials of .01 at an $\alpha = .05$ level 95% of the time when the standard deviation is $\sigma_\varepsilon \leq .05$.

CHAPTER I

INTRODUCTION

If a comparison of two or more systems is to be meaningful it seems reasonable that each system should be operating at its full potential. When the human factor enters into the operation of a system, the system's full potential is not realized until the operator(s) are "fully learned." To determine if the operators are fully learned, one must be able to measure the rate of learning that is occurring over successive periods of operating the system. Assuming learning can be described by a monotonic function, a fully learned status would correspond to a zero rate of learning.

When the crew operating one system is fully learned while the crew operating a second system is not, a negative bias could be introduced into the test results of the second system, rendering an unfair evaluation. This bias which is the result of a difference in the proficiency levels of respective crews on competing systems is a major concern to the U. S. Army Operational Test and Evaluation Agency (OTEA) during operational testing and evaluation of contracted equipment.

Background

This study was prompted by the desire of the U. S. Army Operational Test and Evaluation Agency (OTEA) to determine if a crew or unit is fully learned on the operation of the system being evaluated.

The purpose of operational testing is to provide information for use in an independent evaluation of the military utility, operational effectiveness and suitability of the total system [1]. There are three sequential tests, OTI, OTII, OTIII, characterized by emphasis on testing with typical user operators, crews, or units under realistic conditions. Each sequential test consists of several trials. Data obtained during a particular segment of the sequence is analyzed to determine if the next phase of the test should be conducted or the new system rejected [2].

Operational test I, (OTI), usually is limited in scope and focuses on the primary system function (i.e., firepower of a weapon, mobility of a transport system, etc.). The type of comparison is either against a baseline system or among competing systems. Operational test II, (OTII) is broader in scope and is concerned with testing of engineering prototype equipment and complete test support packages involving entire troop units in controlled field exercises. The comparison is between the new system and the standard system which would be replaced. Operational test III (OTIII), involves evaluating the performance of as large

a unit as feasible, employing the new system versus the same unit employing the current system in use. It is important, therefore, not only to detect if learning is occurring but to detect it early in the OT before additional time and money is expended on obtaining possibly meaningless results. To obtain timely information for deciding to stop the OT will usually require an on site evaluation of the test results. Thus any methodology developed must not only be able to detect learning but must also be applicable in a field environment.

Fundamentals of the Learning Curve

Assuming the performance of the system is dependent on operator proficiency and can be described by a monotonic function, we then have a situation which can be modeled by the basic learning curve function.

Learning defined by improved cycle time or performance over repeated trials can be divided into two distinct phases: threshold learning and conditioned learning. Threshold learning is that learning which occurs prior to the time the operator can do the operation from memory. Conditioned learning is that learning which occurs after the person remembers how to perform the operation without relying on a trial and error procedure. For this research, only the second phase or conditioned learning will be considered.

According to the findings of previous research studies

[3], [5], [8], [9], the learning process can be defined by an equation of the form:

$$z = at^{-b} \tag{1-1}$$

where

z = cycle time

a = a constant which is determined by the cycle time at the beginning of conditioned learning

t = the cycle number from the start of conditioned learning

b = a constant which is determined by the rate of learning over trials.

Although this function is continuous for values of t greater than zero, learning can only be meaningfully evaluated at discrete values of t. This particular equation describes the learning of an operation without any interruption of significant duration which could have a negative effect on previously learned information and skills. The values of the parameters will always be greater or equal to zero.

In conducting trials during a particular phase of the operational test at OTEA, the time between trials sometimes varies but it is believed to have no significant effect on retention of previously gained knowledge and skills.

Recently much interest has been focused on group learning patterns. Several case studies have been conducted to determine if group learning can be described by an equation similar in form to equation (1-1). Although studies

in this area have been somewhat limited in their scope, it appears on the surface that team learning exhibits the same performance curve displayed by individual operators [12].

Another way to examine learning is by a performance curve

$$z = 1-at^{-b} \tag{1-2}$$

This curve is based on the same theory as the "cycle time" curve except the asymptote of this function approaches the value 1 (see Figures 1-1 and 1-2). The performance curve is based on percent achieved from the total possible obtainable. This method of recording learning would be appropriate when accuracy rather than time to completion was the primary objective. The restrictions on parameter "b" are the same as for the learning curve, however, parameter "a" will only take on values in the interval [0,1].

Although the theoretical asymptote for the curve is one when the number of trials approaches infinity, this function can approach any value between zero and one as a working asymptote by using the proper combination of parameters "a" and "b." A working asymptote is referred to here as that value on the curve where the change in performance between trials is so small that it would be considered negligible for practical purposes.

For example, in a given trial of an OT if a weapon is fired at a target 100 times, the total possible performance would be 100 hits or 100 percent. If the weapon is only

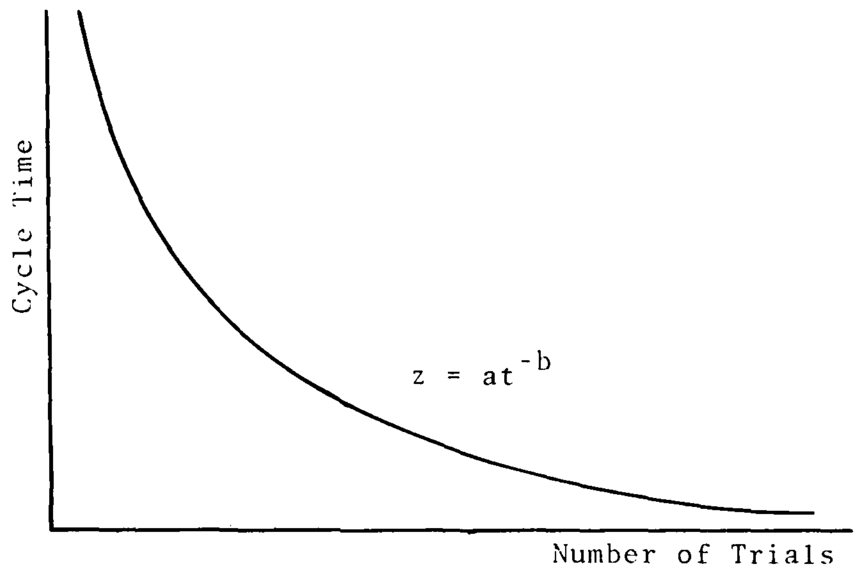


Figure 1-1. Cycle Time Curve

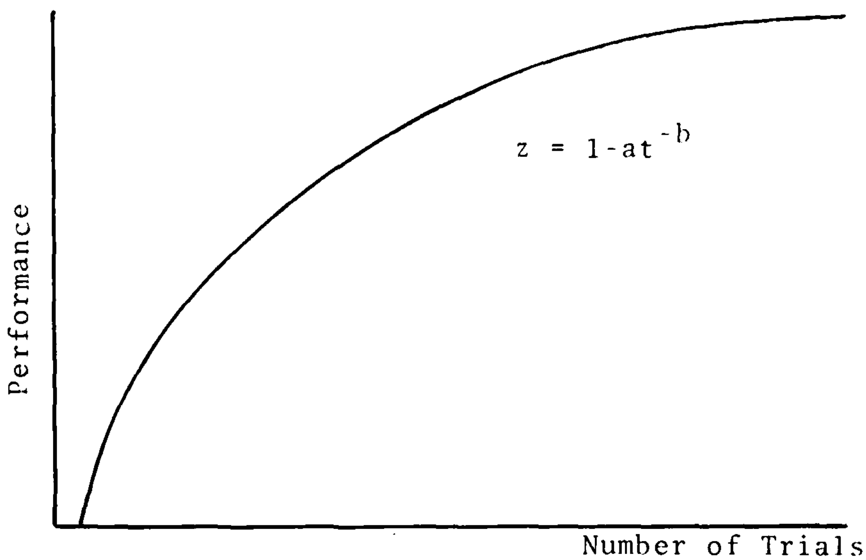


Figure 1-2. Performance Curve

capable of hitting 75 percent of the targets when fired by fully learned operators, then the function describing the learning of the operators over several trials would approach an asymptote of .75 on the performance scale. The theoretical asymptote of the function describing learning would of course still be one; however, for evaluation purposes the operators would be considered fully learned and the weapons capability assumed to be less than 100% accuracy.

In both curves, equation (1-1) and equation (1-2), the value of the function before the first trial and between discrete values of the trials has no significance since there is no measure of knowledge until a trial is completed. Therefore, when dealing with learning curve equations, whether cycle time or performance oriented, the primary concern is the description of learning over discrete values of $t \geq 1$.

Basis for a Performance Curve in Test Results

Before conducting any operational tests at OTEA, the participants undergo a thorough condensed training program on the system to be evaluated. Due to extremely high costs in operating the system, much of the training is conducted piecemeal under simulated conditions which may or may not truly represent the performance of a crew in an actual situation. If a crew's performance could be improved by actually operating the system, then the test results over several trials of an OT segment should reflect improvement

through better scores. If a sufficient number of trials were conducted the test results would eventually level off indicating the crew's performance has peaked and that no further learning is taking place. Thus it seems reasonable, and OTEA test results are currently being evaluated in support of this conclusion, that a performance curve function can be fit to test results of this nature.

Objective

The objectives of this research are two fold. The first objective is to devise a field expedient methodology for testing if the rate of learning is significantly different from zero. The second objective is to develop a methodology which can measure the rate of learning taking place at any particular trial.

General Approach

Since it is well documented that the curve describing performance as learning progresses follows an asymptotic curve, the rate of learning then could be analyzed by either evaluating the first derivative of a curve fit through the data points or examining the slope of a line between data points for specified times. If the first derivative of the curve or the slope of the line is positive, it is an indication that learning is taking place. One procedure to be examined will be to estimate the slope of a linear model fit to the observations and test for its statistical significance.

The test procedure involves simple algebra and requires no information about the parameters of the actual curve.

The test results on the slope using a linear model would only indicate if learning occurred during the trials and can provide no information as to whether the rate of learning was decreasing over the latter trials.

Another drawback using linear methods is that the rate of learning will be tested using an estimate of the variance obtained by fitting a linear model of the form

$$y_i = c + dt_i + \delta_i \tag{1-3}$$

where

y_i represents the observation at trial i

t_i represents trial i

c is the intercept value

d is the slope of the line

δ_i difference between the observation and the line at trial i

when the true model is the nonlinear function

$$y_i = 1-at^{-b} + \varepsilon_i \tag{1-4}$$

The error, ε_i, between any observation, y_i, and its expected value, z_i, is assumed to be independent and normally distributed with expected value of zero and variance

of σ_ε^2, NID $(0, \sigma_\varepsilon^2)$. If the error term such as ε_i is a sum of errors from several sources, then no matter what the probability distribution of the separate errors may be, their sum ε_i will have a distribution that will tend more and more to the normal distribution as the number of components increases by the central limit theorem [13]. An error in a test observation may be a composite of a scoring device error, an error due to a small leak in the system, an error due to an unexpected physical ailment affecting the operator, an error due to changes in wind velocity and so on. The components of this error term would not include those dependent on operator proficiency and likely to decrease with additional repetitions or training. This latter type of error is often used to record learning and would be reflected by the performance curve. The error terms in equation (1-3) may be larger than in equation (1-4) due to a lack of fit of the model which will in turn inflate any estimate of the variance used in testing for the statistical significance of the slope, d. The closer the trial observations are to the asymptote of the expected curve, however, the better the estimate of the variance will be since the lack of fit component will be decreasing. Therefore, a linear method may be appropriate to detect learning if the estimate of the variance is relatively accurate.

Another approach will be to fit a curve to the

observations using a nonlinear regression technique and analyze the location of the trial results in relation to the fitted curve. Although the estimate of the variance using nonlinear techniques will be more accurate than that using the linear estimate, the difficulty in conducting significance testing is that the estimates obtained using nonlinear techniques do not have the linear properties necessary to conduct the known significance tests. It may be possible however to use linear theory results as approximations for determining a confidence region for the parameters of the nonlinear model if the degree of nonlinearity is not too large. If the performance function satisfies this requirement then the rate of learning can be determined by analyzing the approximated confidence interval about the slope at specified trials. This procedure then would provide a means to determine how close to being fully learned the operators are at each trial.

Measure of Nonlinearity

When a model is nonlinear there is an estimation space, however, it is not defined by a set of vectors and may be quite complex. If the estimation space consists of all points with coordinates $\{f(x_1,\theta), f(x_2,\theta), \ldots, f(x_m,\theta)\}$ then minimizing the sum of squares function $ss(\theta)$ corresponds geometrically to finding a point p on the estimation space which is the shortest distance to Y, the vector of observations.

A sample space for a very simple non-linear example involving only n=2 observations y_1 and y_2 taken at $x = x_1$ and $x = x_2$, respectively, and a single parameter θ is illustrated by Draper and Smith [6] and is reproduced in Figure 1-3. The non-linear estimation space consists of the curved line which contains points $\{f(x_1,\theta), f(x_2,\theta)$ as the parameter, θ, varies, and the independent variables x_1, x_2 are fixed. The point Y has coordinates (y_1,y_2) and p is the point of the estimation space closest to Y. When the linearization technique is applied to a non-linear problem, a new origin is selected, say θ_o, and a linearized estimation space in the form of the tangent line at θ_o is then defined. The linear estimation space contains the points $\{f(x_1,\theta_o) + \frac{\partial f(x_1,\theta)}{\partial\theta}$, $f(x_2,\theta_o) + \frac{\partial f(x_2,\theta)}{\partial\theta}$ as θ varies and x_1, x_2 are fixed. However if the rate of change of $f(x,\theta)$ is small at θ_o but increases rapidly, the units on the tangent line may be unrealistic in terms of determining good estimates of the parameters that will minimize the sum of squared errors between the observations and the proposed model. Again Draper and Smith give an excellent illustration of gross inequalities in the systems of units. See Figure 1-4.

In Figure 1-4 the best linear approximation of the true parameter solution from the point $\theta = \theta_o$ is the point $\theta = Q_o$. It is obvious that if the linear solution $\theta = Q_o$ is used as the next starting point on the estimation space we will be further from the best point P then was our original

guess $\theta = \theta_o = 0$. If the degree of non-linearity is not too large, it may be possible to use linear theory results to approximate the confidence region for the nonlinear function. Therefore we need a procedure that will determine when linear theory results provide acceptable approximations to the nonlinear estimation problem.

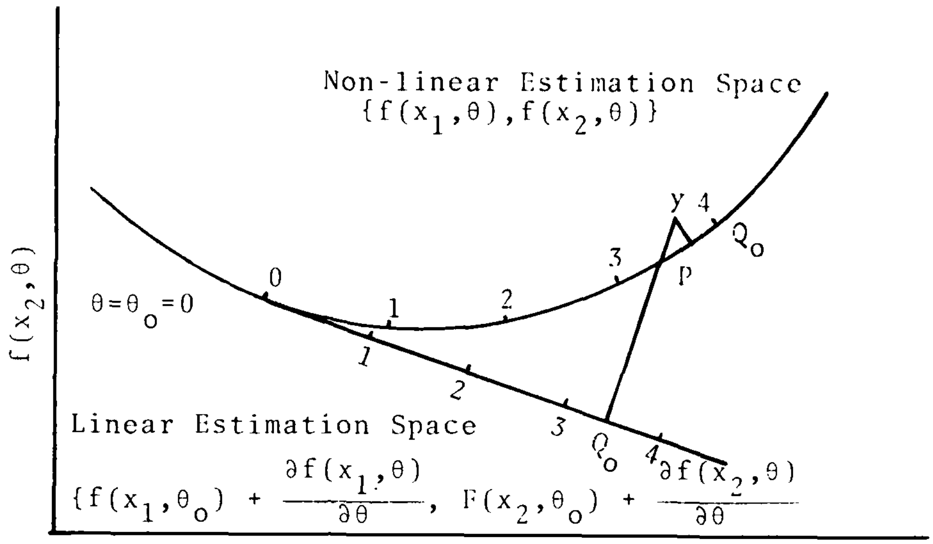


Figure 1-3. Geometric Interpretation of Linearization Method (n=2, θ=1)

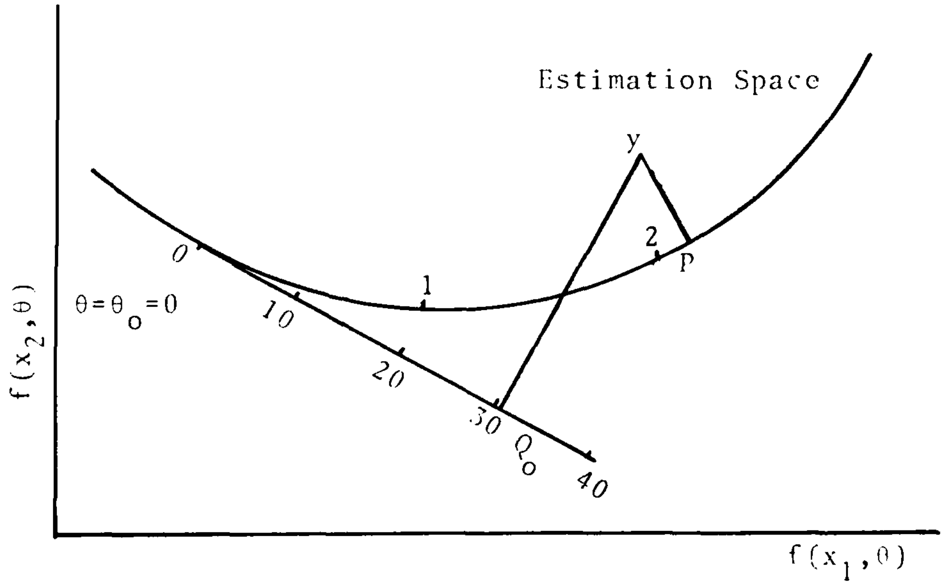


Figure 1-4. Effect on the Linearization Method of Gross Inequalities in the System of Units (n=2, θ=1)

CHAPTER II

METHODOLOGY

The curve describing operator performance on a particular system is well defined in terms of the parameters "a" and "b." This curve can then be used as a basis for determining any given crew's performance level on the system.

Determining the Distribution of the Linear Test Statistic

A method for determining a crew's performance level when the curve is well defined is to examine the expected values of the observations. The procedure would be to find the expected value for each trial observation by minimizing the sum of the squared errors between the observations and the known curve at discrete consecutive trial numbers. If the expected value that corresponds to the last observation is at the asymptote of the curve it is assumed the crew is fully learned. This procedure is summarized as follows.

Given n observations denoted as $y_1, y_2, \ldots, y_n$, label their respective trial numbers k, k+1, ..., k_{n-1}, and find the discrete value of k that will minimize the sum of squared errors between the observations and values on the curve computed at the corresponding observation trial numbers.

The procedure for this is to find the discrete value

of k, k_o, on the known curve that corresponds to the minimum error between the first observation and the value of the known function (see Figure 2-1). Retaining the same labeling for the trial numbers corresponding to the observations, conduct a search over discrete values in the vicinity of k_o, and find the discrete value for k that minimizes the sum of squared errors (Figure 2-2). Find the corresponding trial number, k + N - 1, for the last observation, and then compute value of the known function at this value. If the value of the function at k + N - 1 is the asymptotic value, this corresponds to the situation in which the expected value of the last observation is the asymptote which means a fully learned status.

Another approach to determining the performance level would be to conduct statistical tests on the observations. If the distribution of the error between an observation, y_i, and its expected value, z_i, is NID (0, σ_ε^2), then the distribution of the observations at a given trial is NID (z_i, σ_ε^2). Due to the normality assumption,

$$\frac{y_i - z_i}{\sqrt{\hat{\sigma}_\varepsilon^2}} \tag{2-1}$$

where $\hat{\sigma}_\varepsilon^2$ is an estimate for the true variance

follows a student-t distribution. To test for learning over a series of trials, the test statistic would be

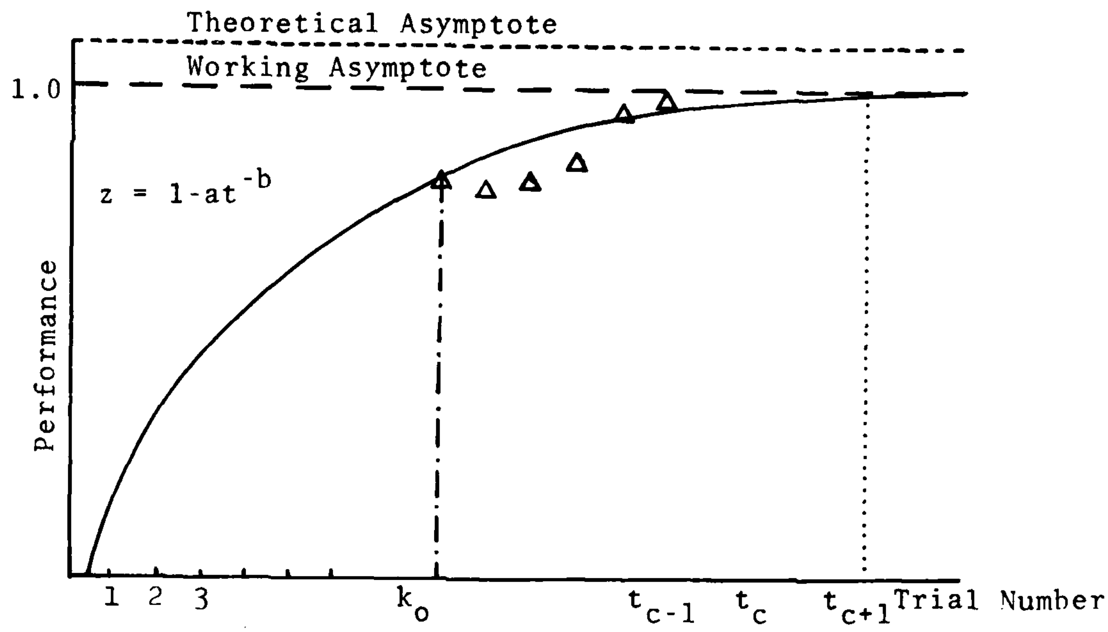


Figure 2-1. Fitting First Observation to the Curve

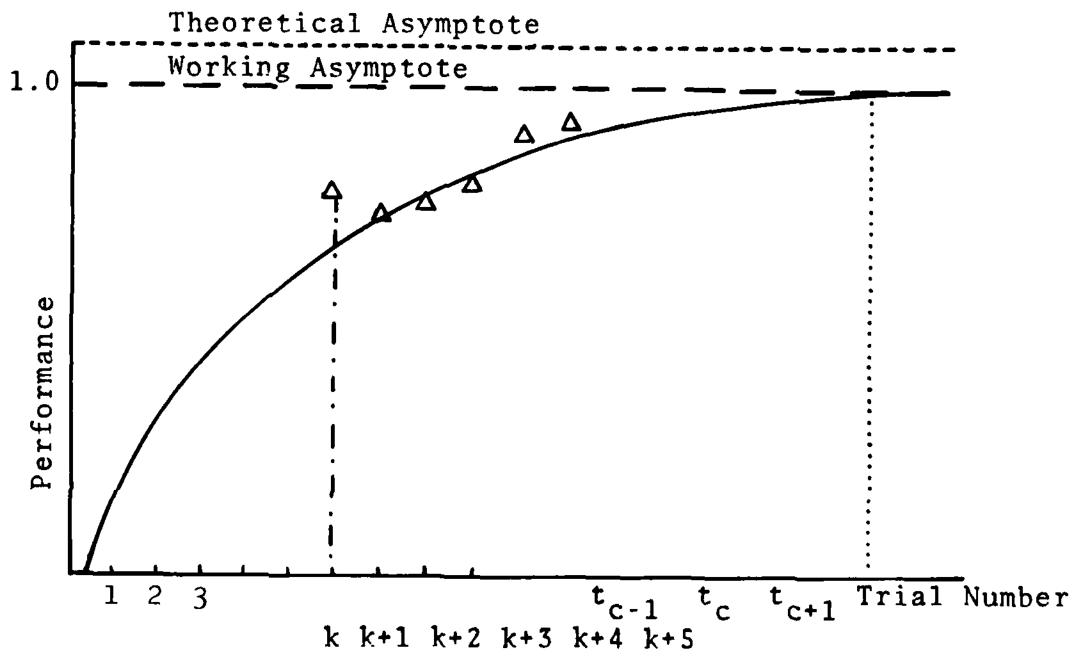


Figure 2-2. Test Observations Fit to the Curve

$$t_o = \frac{\bar{y} - z_c}{\sqrt{\hat{\sigma}_\varepsilon^2}} \tag{2-2}$$

where

z_c = asymptotic value

$\bar{y}$ = average value of the observations

since the expected value for $\bar{y}$ is z_c, when the observations are at the asymptote. If the variance is unknown, then an estimate must be found which will provide an unbiased estimate of the variance in order to conduct the test described in equation (2-2).

This is the basis for conducting the linear tests discussed in Chapter I. If $E(\bar{y}) = z_c$, then all the observations are at the asymptote and the rate of change between observations or the slope of the linear model will be equal to zero. The discussion that follows in the remainder of the section and in the next section will be devoted to obtaining an estimate of σ_ε^2 in order that the t-test may be used to test for learning.

Assuming then that the error term, ε_i, in any observation, y_i, which fits the form

$$y_i = 1 - at^{-b} + \varepsilon_i \tag{2-3}$$

is NID(0, σ_ε^2), a minimum variance unbiased estimate for the variance of an observation about its expected value would be

$$S_y^2 = \frac{\sum_{i=1}^{N} (y_i - z_i)^2}{N} \qquad (2\text{-}4)$$

where y_i is the i^{th} trial observation

z_i is the expected value of y_i

Since in most cases the expected curve of the observations will be unknown, this procedure is of little practical value.

Estimating σ_ϵ^2 When the Paramters Are Unknown

Continuing with the performance curve as defined in equation (1-2), performance will reach an asymptote* as the number of trials increase. There is then some critical trial number, t_c, where further trials will show negligible improvement in proficiency.

Consider a situation where no well defined curve exists. The expected values of the observations are unknown. To estimate the variance using an equation similar in form to equation (2-4), an estimate for z_i must be obtained. Consider the curve in Figure 2-3.

*In the remainder of this research, the word asymptote refers to the value of the curve where the rate of change over future trials is so small (say .00001) that it is considered zero for practical purposes.

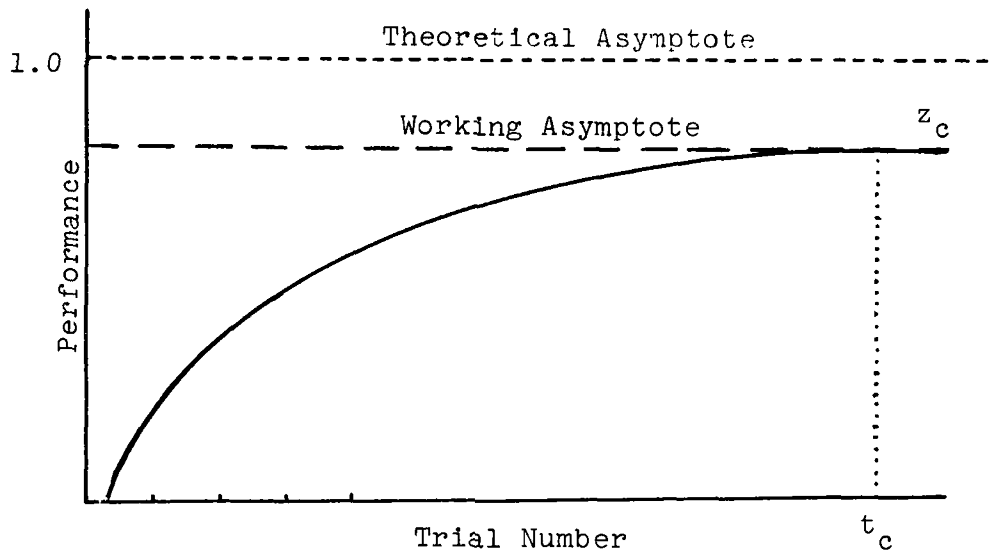


Figure 2-3. Performance Curve with Asymptote at z_c

Assume the results of several trials follow the curve described in Figure 2-3. If the test observations fit the curve beyond trial t_c, the differences in their expected values would be small enough to be considered zero for all practical purposes. The expected value for each observation is then considered to be equal to z_c, the value at the asymptote of the curve. Since the observations in this situation are all distributed NID (z_c, σ_ε^2), an unbiased estimate for z_i in equation (2-4) would be the average of the observations.

$$\bar{y} = \sum_{i=1}^{N} y_i/N \qquad (2\text{-}5)$$

where $y_i = z_c + \varepsilon_i$

$$E(\bar{y}) = E[\sum_{i=1}^{N} (z_i+\varepsilon_i)/N]$$

$$E(\bar{y}) = \frac{\sum_{i=1}^{N} E(z_i+\varepsilon_i)}{N}$$

$$E(\bar{y}) = \frac{\sum_{i=1}^{N} z_i}{N} \qquad \text{but } z_i = z_c$$

therefore $$E(\bar{y}) = N\, z_c/N = z_c$$

Thus an efficient unbiased estimate for the variance [11], using the sample average as the estimate of the z_i's when *the observations* are all at the asymptote is

$$S_\varepsilon^2 = \frac{\sum_{i=1}^{N} (y_i-\bar{y})^2}{N-1}$$

When the observations are not at the asymptote, this estimate of the variance will be inflated since the average of the observations, $\bar{y}$, will no longer be an unbiased estimate for the expected value of each observation. The amount of bias will be a function of the distance that the expected values of the observations are from the asymptote of the true curve.

Since the performance curve function is non-linear, there is no easy-to-apply procedure to obtain an unbiased estimate for the variance when the parameter

values are unknown.

Several alternative techniques for estimating the variance will be examined in terms of the expected bias to determine a minimum biased estimator. The deviation of several estimators is contained in Appendix B.

The three estimators to be investigated for estimating σ_ε^2 are:

$$(OBS)^2 = \sum_{i=1}^{N} (y_i - \bar{y})^2 / N-1 \tag{2-7}$$

$$(SEX)^2 = (N-1) \sum_{i=1}^{N-1} (x_i - \bar{x})^2 / 2N(N-2) \tag{2-8}$$

$$(SER)^2 = (N^2+1)(N+2)(MS_E)/(N^3-2N^2+N+1) \tag{2-9}$$

In summary, the expected values for the respective variance estimators are:

$$E\{\frac{(N^2+1)(N-2)MS_E}{N^3-2N^2+N+1}\} = \sigma_\varepsilon^2 + \frac{12}{N(N^3-2N^2+N+1)} [\frac{N^3+N}{12} \sum_{i=1}^{N} z_i^2 - \frac{(N^2+1)(\sum_{i=1}^{N} z_i)^2}{12} - (\sum_{i=1}^{N} t_i z_i - \frac{N+1}{2} \sum_{i=1}^{N} z_i)^2]$$

$$E\{\frac{(N-1)\sum_{i=1}^{N-1}(x_i-\bar{x})^2}{2N(N-2)}\} = \sigma_\varepsilon^2 + \frac{N-1}{2N(N-2)} [\sum_{i=1}^{N-1} (z_{i+1}-z_i)^2 - \frac{(z_N-z_1)^2}{N-1}]$$

$$E\{\frac{\sum_{i=1}^{N-1}(y_i-\bar{y})^2}{N-1}\} = \sigma_\varepsilon^2 + \frac{1}{N-1} [\sum_{i=1}^{N-1}(z_i-\bar{z})^2]$$

Note that if the expected values of the observations, the z_i's, were all at the asymptote, the bias factor in all three equations is zero. As the difference between the expected values of the observations and the asymptotic value of the curve increases however, the expected bias value associated with each procedure changes. The relative usefulness of the estimators will be numerically analyzed in Chapter III.

Variance of the Estimators of σ_ε^2

If the variances of the estimators of σ_ε^2 are not *significantly different*, then the expected bias may be the only criteria necessary to determine the best estimator (see Figure 2-4a). If on the other hand the variances of the estimators differ significantly, then it is possible that the minimum biased estimator is not the best estimator in terms of the percent of estimates within the specified tolerance limits (see Figure 2-4b).

When the allowable error tolerance for the variance estimate is δ, then estimator 1, in Figure (2-4b), is the better estimator. When the allowable error is 2δ, however, then estimator 2 is better since it has the largest percent of its estimates within the tolerance limits.

The expected bias associated with each estimator is a function of the expected values of the observations and is

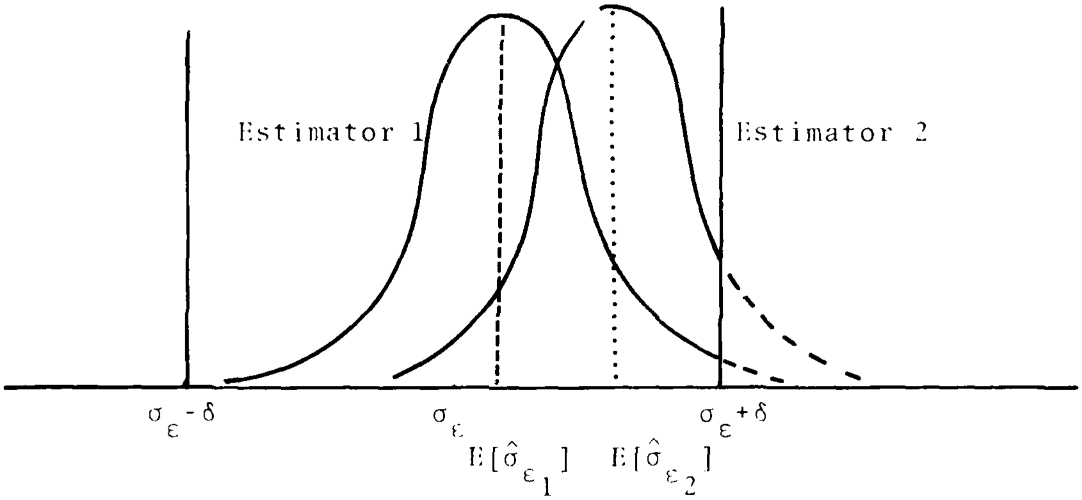


Figure 2-4a. Possible Distribution of Estimates when the Variances of the Estimates are Equal

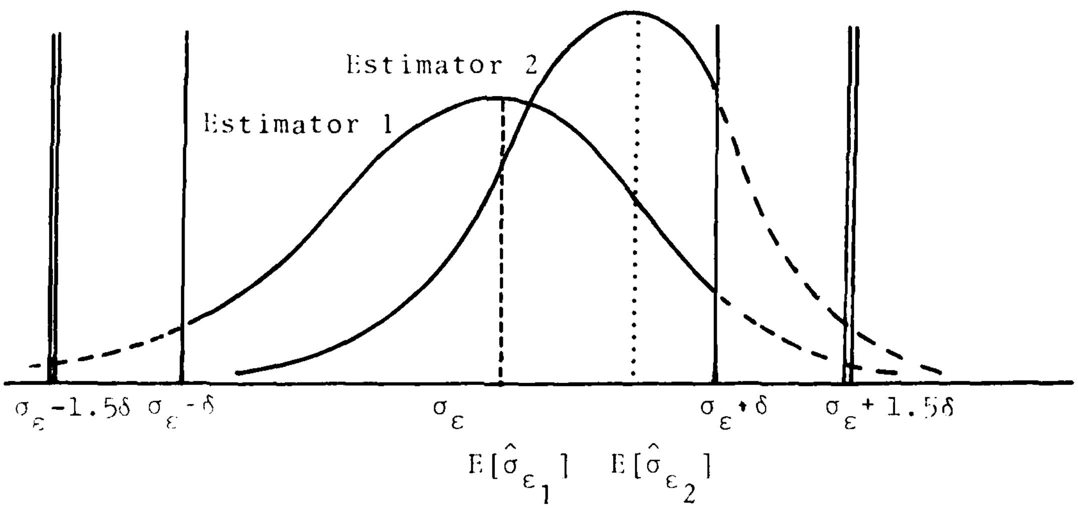


Figure 2-4b. Possible Distribution of Estimates when the Variance of the Estimators are not Equal

only affected by changes in the parameter values or the sample size. Therefore, the difference between the expected value of the estimate for σ_{ε}^2 and the value of σ_{ε}^2 for specified a, b, and N, will not change as the size of σ_{ε}^2 changes.

The magnitude of σ_{ε}^2 will be a factor however, in determining the expected bias in the estimate of σ_{ε}. See Figures (2-5) and (2-6). This is important since this is the value that will be used to compute any statistical test on the slope. As the process variance increases for specified parameter values and sample size, the expected value of the estimate of σ_{ε} approaches the true value of σ_{ε}. As σ_{ε}^2 increases though, there will be a corresponding increase in the variance of the estimator. This will affect the dispersion of the estimates and thus could also affect the solution of an estimator. See Figure (2-5) and Figure (2-6).

If the contribution of σ_{ε}^2 to the estimator variance for one estimator is larger than for another, a situation could also occur as depicted in Figure (2-7a) and Figure (2-7b).

When σ_{ε}^2 is small as in Figure (2-7a), estimator 2 is better, however, in Figure (2-7b) where σ_{ε}^2 is large, estimator 1 appears to do as well or better than estimator 2. To determine a best estimator then, the effect of the variance on the estimates as well as that of the parameter values and sample size must be examined.

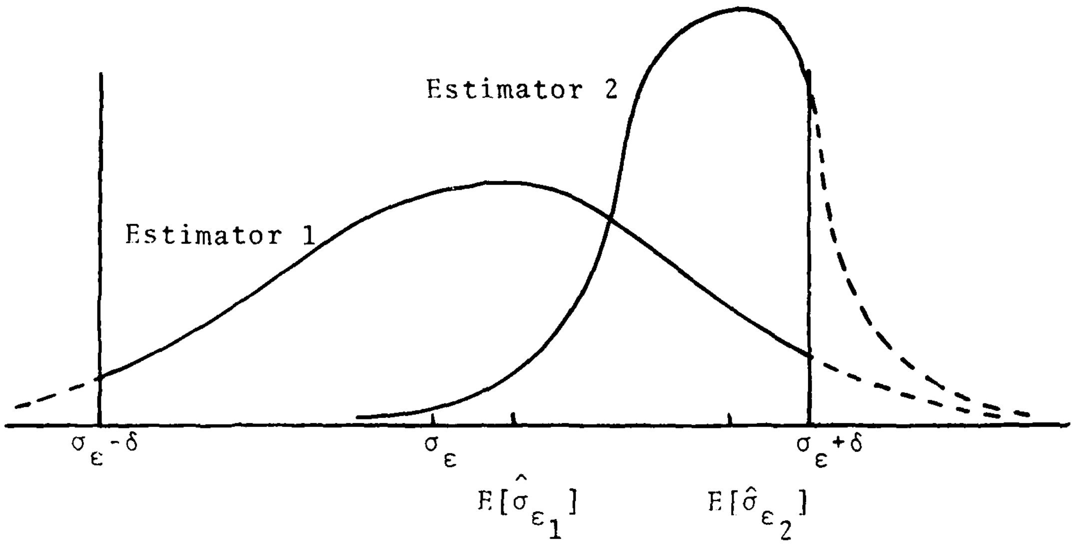


Figure 2-5. Expected Bias Large and Dispersion of Estimates Small When σ_ε is Small

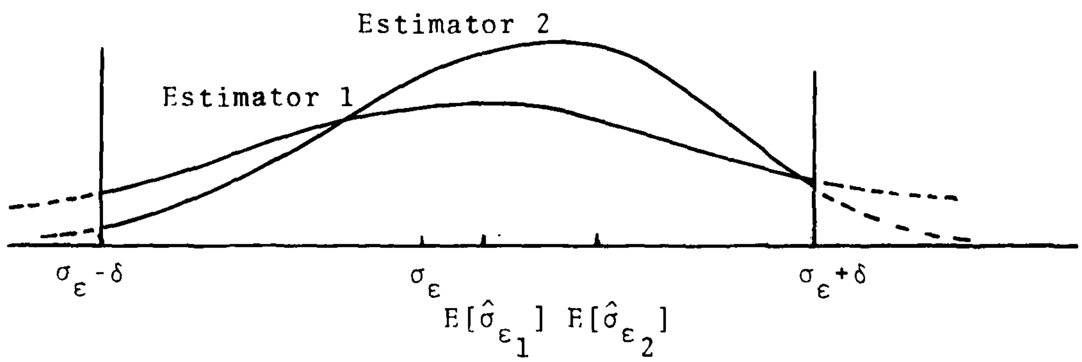


Figure 2-6. Expected Bias Small and Dispersion of Estimates Large When σ_ε is Large

An evaluation of the variance of the estimator as well as the expected bias factor will thus be required before selecting the best estimator of σ_ε^2. An effort was made to obtain a closed form expression for the variance of each estimator in terms of the true variance, σ_ε^2, and the expected values of the observations, the z_i's. Due to the complex forms involved, this approach was abandoned in favor of analysis by computer simulation. The results of the simulation study will be presented in Chapter III.

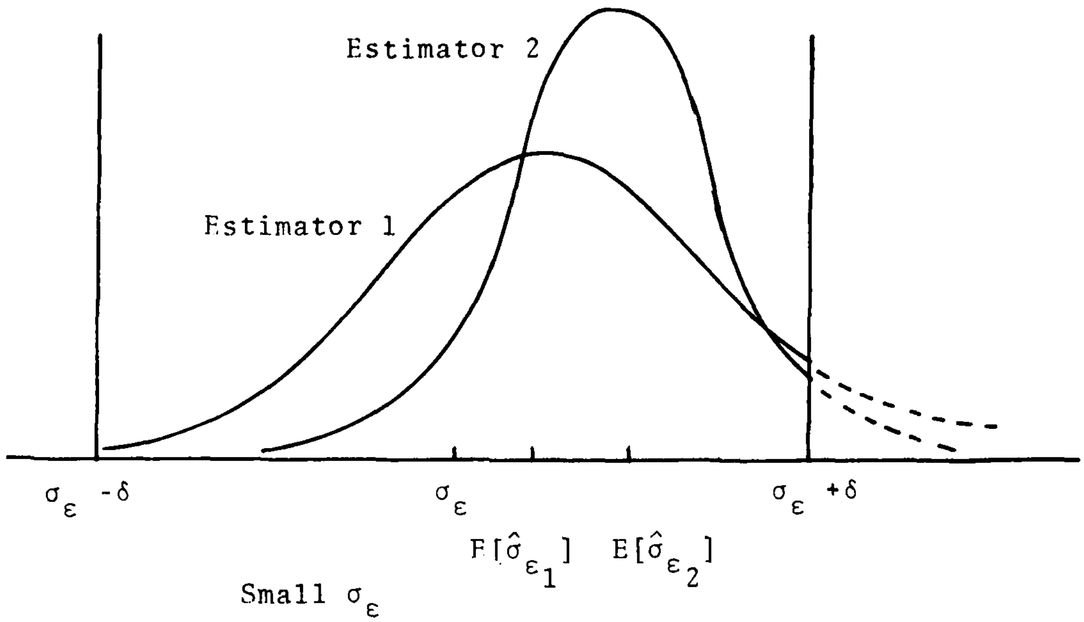


Figure 2-7a. Possible Effect on Dispersion of the Estimates of σ_ε as the True Value of σ_ε Increases

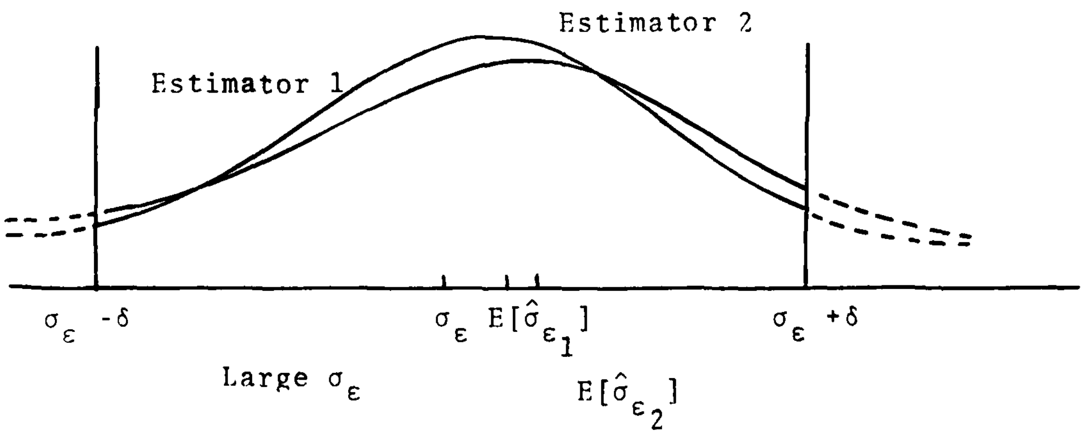


Figure 2-7b. Possible Effect on Dispersion of the Estimates of σ_ε as the True Value of σ_ε Increases

Linear Methods to Test for Learning

Since the performance curve describing the progress of training is an asymptotic function, it may be possible to determine when a unit is fully learned by examining the slope of a line fit through the trial observations or examin- the rate of change between the observations. If the rate of change is large, then learning is taking place and when the rate of change approaches zero, a very small amount of learning is occurring which corresponds to approaching a "fully learned" status. An appropriate method to determine the level of performance then would be to test if the slope of the linear model fit through the observations or the rate of change between observations is significantly greater than zero.

Linear Methods of Analysis

As discussed earlier, if the rate of learning is large and the error variance of reasonable size, it may be possible to detect learning by examining the slope using linear approximation methods. Several methods, which can be solved by simple hand calculations, are examined using the variance of the slope as a basis for comparison in selecting which is best. Those linear approximation methods which provide the smallest variance for the slope will be selected for further study and possible application.

a. "Average of Two Groups (ATG)". In order to obtain an estimate of the slope it is necessary to have at least two points. The average of two groups method divides the data at the midpoint of the trials into two groups. The average of the observations is computed for each group, and the difference between the averages tested to determine if this difference is significantly greater than zero. If there were an odd number of observations, the middle observation would not be considered in the computations.

Let D_E = average of observations in first group

D_L = average of observations in second group

$$d_{ATG} = D_L - D_E \tag{2-10}$$

then

$$Var(d_{ATG}) = Var(D_L - D_E)$$

$$Var(d_{ATG}) = Var(\sum_{i=1}^{N/2} y_i/(N/2)) + Var(\sum_{i=N/2+1}^{N} y_i/(N/2))$$

$$Var(d_{ATG}) = \frac{(N/2)\sigma_\varepsilon^2}{(N/2)^2} + \frac{(N/2)\sigma_\varepsilon^2}{(N/2)^2}$$

$$Var(d_{ATG}) = \frac{4\sigma_\varepsilon^2}{N} \tag{2-11}$$

b. "Average of Consecutive Differences (ACD)". In this method the average of the differences between consecutive observations is analyzed to determine if learning is occurring.

$$d_{ACD} = \sum_{i=1}^{N-1} x_i/N-1 \qquad (2\text{-}12)$$

$$Var(d_{ACD}) = Var(\sum_{i=1}^{N-1} x_i/N-1)$$

$$Var(d_{ACD}) = \frac{1}{(N-1)^2} Var\ (y_n - y_1)$$

$$Var(d_{ACD}) = 2\sigma_\varepsilon^{\ 2}/(N-1)^2 \qquad (2\text{-}13)$$

c. <u>"Average of Two Groups Using Consecutive Differences (ATGCD)"</u>. This method is a combination of methods "a" and "b". The observations are divided into two groups and the differences between observations are computed in each group. Then computing the difference between the average obtained in the latter group with that of the earlier group one obtains an estimate of the slope with corresponding variance as follows:

$$D_E = \frac{\sum_{i=1}^{N/2} (y_{i+1} - y_i)}{N/2}$$

$$D_L = \frac{\sum_{N/(2+1)}^{N} (y_{i+1} - y_i)}{N/2}$$

$$d_{ATGCD} = D_L - D_E \quad . \qquad (2\text{-}14)$$

$$Var(d_{ATGCD}) = Var(D_L - D_E) = \frac{16\sigma_\varepsilon^{\ 2}}{(N-2)^2} \qquad (2\text{-}15)$$

As in the ATG method, if there is an odd number of observations, the middle observation would be disgarded.

d. "Linear Least Squares Regression (LLSR)". This method fits a line through the observations such that the total of the distances squared of the observations from the line is minimized. Draper and Smith [6] provide an excellent description of the details involved. A brief summary of the procedure is given later in this Chapter.

An estimate of the slope is defined as:

$$d_{LLSR} = \frac{\sum_{i=1}^{N} (t_i-\bar{t})(z_i)}{\sum_{i=1}^{N} (t_i-\bar{t})^2} \tag{2-16}$$

where t_i = trial number i

z_i = expected value of the observation at time i

and the slope variance is defined as

$$\mathrm{Var}(d_{LLSR}) = \frac{\sigma_\varepsilon^{\,2}}{\sum_{i=1}^{N}(t_i-\bar{t})^2} = \frac{\sigma_\varepsilon^{\,2}}{\sum_{i=1}^{N} t_i^{\,2} - 2\bar{t}\Sigma t_i + N\bar{t}^2} \tag{2-17}$$

Since the trials are consecutive from 1 to N the following closed form expressions for t can be used in the above equation

$$\sum_{i=1}^{N} t_i^{\,2} = \frac{N(N+1)(2N+1)}{6} \tag{2-18}$$

$$\bar{t} = \frac{N+1}{2}$$

$$t_i = \frac{N(N+1)}{2}$$

Then

$$\mathrm{Var}(d_{LLSR}) = \frac{\sigma_\varepsilon^2}{\frac{N(N+1)(2N+1)}{6} - 2\left(\frac{N+1}{2}\right)\left(\frac{N(N+1)}{2}\right) + N\left(\frac{N+1}{2}\right)^2}$$

$$\mathrm{Var}(d_{LLSR}) = \frac{12\sigma_\varepsilon^2}{N(N^2+1)} \qquad (2\text{-}19)$$

For N>2, the linear last squares regression method provides the best estimates for the variance of the slope. See Figure (2-8). The average of consecutive differences, ACD, method was the next best procedure. The average of two groups method becomes the third best procedure when N=8, but the estimate of the variance of the slope is still quite large in comparison to the ACD and the LLSR methods.

From the several methods considered for detecting learning through analyzation of the slope, the two best procedures, using the minimum variance as the selection criteria, appears to be the average of consecutive differences (ACD) method and the linear least squares regression (LLSR) method.

To complete the analysis of the linear approximation procedures, the expected value of the estimate of the slope using the ACD method and the LLSR method will be examined. Let ℓ represent the true average rate of learning for the process over N trials. An expression for ℓ then is

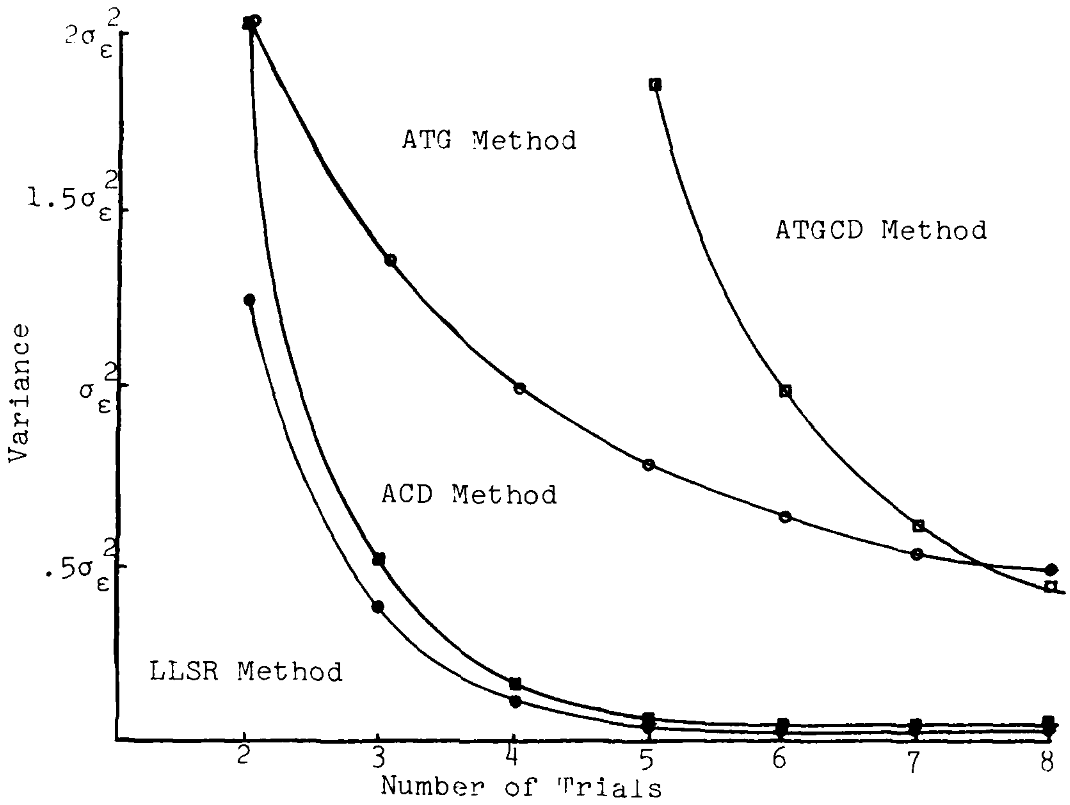


Figure (2-8). Variance of the Average Slope Estimates

$$\ell = \frac{z_N - z_1}{N-1} \tag{2-20}$$

$$\ell = \frac{a - at_N^{-b}}{N-1}$$

The expected value of an estimate of ℓ using the ACD method is

$$E[\hat{d}_{ACD}] = \frac{\sum_{i=1}^{N}(y_{i+1} - y_i)}{N-1}$$

$$E[\hat{d}_{ACD}] = E[\frac{y_N - y_1}{N-1}]$$

$$= E\left[\frac{(z_N+\varepsilon_N)-(z_1+\varepsilon_1)}{(N-1)}\right]$$

$$= \frac{(1-at_N{}^{-b}+\varepsilon_N)-(1-at_1{}^{-b}+\varepsilon_1)}{N-1}$$

$$E[\hat{d}_{ACD}] = \frac{a-at_N{}^{-b}}{N-1} \text{ since } t_1 = 1 \text{ and } E(\varepsilon_i) = 0 \qquad (2\text{-}21)$$

As expected, the estimate of the slope using the ACD method is unbiased.

The expected value of an estimate of ℓ using the LLSR method is derived as follows:

$$E[\hat{d}_{LLSR}] = E\,\frac{\sum_{i=1}^{N}(t_i-\bar{t})y_i}{\sum_{i=1}^{N}(t_i-\bar{t})^2}$$

$$= \frac{1}{\sum_{i=1}^{N}(t_i-\bar{t})^2}\, E[(t_1-\bar{t})(z_1+\varepsilon_1)+(t_2-\bar{t})(z_2+\varepsilon_2) + \ldots$$

$$+ (t_N-\bar{t})(z_N+\varepsilon_N)]$$

$$= \frac{1}{\sum_{i=1}^{N}(t_i-\bar{t})^2}\,[(t_i-\bar{t})z_1+(t_2-\bar{t})(z_2 + \ldots + (t_N-t)z_N]$$

Replacing the $\sum_{i=1}^{N} t_i$'s with expressions in terms of N we get

$$= \frac{12}{N(N^2+1)} [(t_1 - \frac{N+1}{2})z_1 + (t_2 - \frac{N+1}{2})z_2 + \ldots$$

$$+ (t_N - \frac{N+1}{2})z_N]$$

$$= \frac{12}{N(N^2+1)} \sum_{i=1}^{N} t_i z_i - \frac{6(N+1)}{N(N^2+1)} \sum_{i=1}^{N} z_i$$

Substituting $1-at_i^{-b}$ for z_i we get an expression of the form:

$$= \frac{12}{N(N^2+1)} \sum_{i=1}^{N} t_i(1-at_i^{-b}) - \frac{6(N+1)}{N(N^2+1)} \sum_{i=1}^{N} (1-at_i^{-b})$$

$$= \frac{12}{N(N^2+1)} [\sum_{i=1}^{N} t_i - a \sum_{i=1}^{N} t_i^{-b+1}] - \frac{6N(N+1)}{N(N^2+1)} + \frac{6(N+1)}{N(N^2+1)} a \sum_{i=1}^{N} t_i^{-b}$$

$$E[\hat{d}_{LLSR}] = \frac{12a}{N(N^2+1)} [\frac{(N+1) \sum_{i=1}^{N} t_i^{-b}}{2} - \sum_{i=1}^{N} t_i^{-b+1}] \qquad (2-23)$$

Subtracting the true value of the average rate of learning from the estimate in equation (2-23) we can obtain the expected bias.

Let β represent the amount of bias in equation (2-23), then

$$\beta = \frac{12a}{N(N^2+1)} [\frac{(N+1) \sum_{i=1}^{N} t_i^{-b}}{2} - \sum_{i=1}^{N} t_i^{-b+1}] - [\frac{a-at_N^{-b}}{N-1}] \qquad (2-24)$$

When the value of b equals zero in equation (2-24) which corresponds to the expected values of the observations all being equal to the asymptotic value, the expected bias is zero. The bias factor in equation (2-24) is negative which means we should expect the estimate for the average rate of learning using the LLS method to be less than the value defined in equation (2-21).

In summary, we now have two methods to test for learning which exhibit the following characteristics:

(i) The LLSR method provides a biased estimate for the average rate of learning; however, it has the smallest slope variance of all methods examined.

(ii) The ACD method provides an unbiased estimate for the average rate of learning but, the slope variance is larger than that in the LLSR procedure.

The bias factor in equation (2-24) will be computed and its effect analyzed for various combinations of parameter values and sample size in Chapter III. A discussion of the two best linear test procedures is in the next two sections.

Average of Consecutive Differences (ACD) Method

If the test observations are at the asymptote of the curve, the difference between the observations will follow a normal distribution with mean of zero and variance of $2\sigma_{\varepsilon}^{2}$. Given N trial observations with no information concerning the

true variance or the parameters of the actual curve, we examine the average of the consecutive differences in the observations which corresponds to the average rate of learning over the N trials.

Define x_i as the difference between observation y_{i+1} and y_i. Then

$$E(\bar{x}) = E[\sum_{i=1}^{N-1} x_i/N-1]$$

$$E(\bar{x}) = E[\sum_{i=1}^{N-1} (y_{i+1}-y_i)/N-1]$$

$$E(\bar{x}) = \frac{1}{N-1}\sum_{i=1}^{N-1} E(y_{i+1}-y_i)$$

$$E(\bar{x}) = \frac{1}{N-1}\sum_{i=1}^{N-1} (z_c-z_c)$$

z_c = expected value of an observation at the asymptote

$$E(\bar{x}) = u_x = 0$$

From equation (2-13) the variance of the average of the consecutive differences was defined as:

$$\text{Var } \bar{x} = 2\sigma_\varepsilon^2/(N-1)^2$$

where $\bar{x} = \hat{d}_{ACD}$

Using an unbiased estimate of σ_ε^2, the test statistic for the differences will follow the t-distribution. A test to determine if the average slope between observations is significantly different from zero would be:

H_o: slope ≤ 0 fully learned

H_i: slope > 0 not fully learned

Compute:

$$t_o = \frac{\bar{x} - u_x}{2S_y^2/(N-1)^2} \tag{2-25}$$

Testing at a significance level of α,

If $t_o > t_{\alpha, N-1}$ Reject H_o

If $t_o \leq t_{\alpha, N-1}$ Do not reject H_o

Since it is quite possible that the trial observations will not all be at the asymptote, our estimate of the variance will be inflated due to bias as discussed earlier. Using a biased estimate of the variance will reduce the power of the test which is defined as:

Power of test = 1.0-P [failing to reject H_o when H_o is false]

The rationale for this interpretation is that using an inflated estimate for the slope variance in equation (2-25) increases the probability of failing to reject the null hypothesis, H_o, when it is not true. This is because if the test statistic

$$t_o = \frac{\bar{x} - o}{\sqrt{2 Est(\sigma_\varepsilon^2)} / (N-1)}$$

is greater than $t_{\alpha, N-2}$ using a biased estimate for the variance in the denominator, it will remain greater than $t_{\alpha, N-2}$ when an unbiased estimate is used. On the other hand, if t_o is greater than $t_{\alpha, N-2}$ using an unbiased estimate for the variance, it may not be greater if a biased estimate is used, thus creating a situation where we would fail to reject H_o when in fact it is false. Therefore if the test statistic indicates the slope is significantly greater than zero, we can assume learning is occurring, otherwise the results may be unreliable.

A summary of the computations required to conduct the test on the slope using the ACD method follows:

Compute an estimate of the slope:

$$\hat{d}_{ACD} = \frac{y_N - y_1}{N-1} \qquad (2\text{-}26)$$

Compute an estimate for the standard error of the slope:

$$\text{s.e.}(\hat{d}_{ACD}) = \frac{2[\text{Est}(\sigma_{\varepsilon}^{2})]}{(N-1)^2} \quad (2\text{-}27)$$

Compute the test statistic:

$$t_0 = \frac{\hat{d}_{ACD}-0}{\text{s.e.}(\hat{d}_{ACD})} \quad (2\text{-}28)$$

Compare test statistic at desired significance level

If $t_0 > t_{\alpha,N-2}$ Learning is occurring, reject H.

If $t_0 \leq t_{\alpha,N-2}$ No conclusion

The Linear Least Squares Regression (LLSR) Method

As discussed earlier, this method involves fitting a linear model of the form

$$y_i = c + d_{LLSR}t_i + \delta_i \quad (2\text{-}29)$$

where y_i = value of the observation

t_i = trial number that corresponds to y_i

c = y intercept value

d_{LLSR} = difference between the observation and the line at trial i.

δ_i = difference between the observation and the line at trial i.

through the observations, based on minimizing the sum of squares of deviations of the observations from this line.

The slope of the line, d_{LLSR}, at the asymptote of the

curve is essentially the average of the differences between the expected values of successive observations. Therefore the test statistic for the slope of the line using the LLSR method follows the same distribution as the test statistic for the differences in the ACD method, namely the t-distribution the test for learning would then be:

H_o: slope ≤ 0 Fully learned

H_1: slope > 0 Not fully learned

Compute

$$t_o = \frac{\hat{d}_{LLSR} - 0}{\sqrt{\frac{12\ Sy^2}{N(N^2+1)}}} \tag{2-30}$$

where $\hat{d}_{LLSR}$ is the estimated slope for the linear model.

Compare at α significance level

If $t_o > t_{\alpha, N-2}$ Learning is occurring, Reject H_0

If $t_o \leq t_{\alpha, N-2}$ No conclusion

As the bias factor increases, the width of the confidence interval for the slope increases or expressed another way, the test statistic, t_o, decreases.

As was previously shown, using a biased estimate for the variance versus an unbiased estimate when testing the slope, equation (2-28), results in a less powerful test.

This also holds true for the estimate of the average rate of learning. As was the case with the ACD method, the test results may be unreliable if the slope test statistic is less than the comparison value, $t_{\alpha,N-2}$.

In the discussion of the method that follows, it is assumed that the reader is familiar with the linear least squares regression procedure. The LSSR equations required to conduct the slope test are presented without the derivation. (For a development of the equations, see Draper and Smith [6].)

The slope of the line is found by minimizing the sum of the squared errors between the observations and the fitted line equation (2-29). By taking the first derivative of the equation for the sum of squared errors with respect to each of the parameters c and d_{LLSR}, it is possible to solve the two resulting equations simultaneously to obtain an estimate for the slope.

$$d_{LLSR} = \frac{\sum_{i=1}^{N} (t_i - \bar{t}) y_i}{\sum_{i=1}^{N} (t_i - \bar{t})^2} \tag{2-31}$$

where

t_i represents trial number i

y_i represents the observation that corresponds to trial i

$\bar{t}$ is the average of the t_i

$\bar{y}$ is the average of the y_i

with this estimate of the slope and replacing S_y^2 in equation (2-30) with an estimate for σ_ε^2 we can conduct a test for learning. Using an appropriate estimate for σ_ε^2, the estimate of the standard error for $\hat{d}_{LLSR}$ is

$$s.e.(\hat{d}_{LLSR}) = 12\ Est(\sigma_\varepsilon^2)/[N(N^2+1)]$$

In summary, to use the LLSR method to test for learning, one need only compute the following steps

$$\hat{d}_{LLSR} = \frac{\sum_{i=1}^{N} (t_i-\bar{t})(y_i)}{\sum_{i=1}^{N} (t_i-\bar{t})^2} \tag{2-32}$$

$$s.e.(\hat{d}_{LLSR}) = \frac{12\ Est\ (\sigma_\varepsilon^2)}{N(N^2+1)} \tag{2-33}$$

Test procedure:

$$H_0:\ \hat{d}_{LLSR} \leq 0$$
$$H_1:\ \hat{d}_{LLSR} > 0$$

Compute:

$$t_0 = \frac{\hat{d}_{LLSR}-0}{s.e.(\hat{d})} \tag{2-34}$$

If $t_o > t_{\alpha, N-2}$ reject H_o otherwise the test must be regarded as inconclusive.

Non-Linear Method to Test for Learning

If the degree of non-linearity in a particular performance curve is small enough, it may be possible to examine the slope based on linear techniques. Measures of non-linearity have been developed by Beale [4] which indicate when the degree of non-linearity in a non-linear function is small enough to justify approximations using linear theory results. A review of Beales procedure follows.

Consider the non-linear model $\eta = f(\bar{x},\bar{\theta})$ where $\bar{\theta}$ is a (px1) vector of parameters and $\bar{x}$ is a vector of independent variables. Given n independent observations on the response $\eta = (y_1, y_2, \ldots, y_n)$, a least squares estimate of the parameters $\bar{\hat{\theta}} = (\theta_1, \ldots, \theta_p)$ is obtained. Then the tangent plane approximation to the solution locus, estimation space, in the neighborhood of $\bar{\hat{\theta}}$ is given by:

$$\tau_i(\bar{\theta}) = \eta_i(\bar{\hat{\theta}}) + \sum_{j=1}^{p} (\theta_j - \hat{\theta}_j) \left. \frac{\partial f(x_i, \bar{\theta})}{\partial \theta_i} \right|_{\hat{\theta}} \quad i = 1,2,\ldots,n \quad (2\text{-}35)$$

or

$$\bar{\tau}(\theta) = \bar{\eta}(\bar{\hat{\theta}}) + x(\bar{\theta} - \bar{\hat{\theta}})$$

where $\eta_i(\bar{\hat{\theta}}) = f(x_i, \hat{\theta})$. Since $\bar{\tau}(\bar{\theta})$ differs from the actual point $\bar{\eta}(\bar{\theta})$ because of the non-linearity of equation [1-2], a

crude measure of non-linearity would be

$$Q_\theta = \sum_{w=1}^{m} \sum_{i=1}^{n} [\eta_i(\bar{\theta}_w) - (\eta_i(\hat{\bar{\theta}}) - \sum_{j=1}^{p} (\theta_{jw} - \hat{\theta}_j) \frac{\partial f(x_i, \bar{\theta})}{2\theta j} |_{\hat{\bar{\theta}}})]^2 \tag{2-36}$$

$$Q_\theta = \sum_{w=1}^{m} ||\bar{\eta}(\bar{\theta}_w) - \bar{\tau}(\bar{\theta}_w)||^2$$

Q_θ is defined by Guttman and Meeter [6] as the sum of squares of the distances (in sample space) from the points $\bar{\eta}(\bar{\theta}_w)$ to the associated points $\bar{\tau}(\bar{\theta}_w)$ on the tangent plane. By dividing Q_θ by the quantity

$$\sum_{w=1}^{m} \{ \sum_{i=1}^{n} [\eta_i(\bar{\theta}_w) - \eta_i(\hat{\bar{\theta}})]^2 \}^2 = \sum_{w=1}^{m} ||\bar{\eta}(\bar{\theta}_w) - \bar{\eta}(\hat{\bar{\theta}})||^4 \tag{2-37}$$

the sum of squared distances is normalized. Guttman and Meeter [6] go on to explain that since Q_θ has the dimension of the square of an observation and the quantity in equation (2-37), the dimension of the fourth power of an observation, then the quantity

$$\hat{N}_\theta = ps^2 \sum_{w=1}^{m} ||\bar{\eta}(\bar{\theta}_w) - \bar{\tau}(\bar{\theta}_w)||^2 / \sum_{w=1}^{m} ||\bar{\eta}(\bar{\theta}_w) - \bar{\eta}(\theta)||^4$$

where s^2 is an estimate of σ^2 (the variance of the observations), is a dimensionless quantity. This value of $\hat{N}_\theta$ can be regarded as the estimated normalized measure of the

non-linearity of the model when expressed in terms of the parameters $\bar{\theta}$. Beale says that the linear approximation is satisfactory if

$$\hat{N}_{\theta} \leq 0.01/F_{\alpha,p,N-p} \qquad (2\text{-}39)$$

since the root mean square value of the discrepancy vector $\bar{\eta}(\bar{\theta}_w)-\bar{\tau}(\bar{\theta}_w)$ is less than one-tenth the length of the intended vector $\bar{\tau}(\bar{\theta}_w)-\bar{\eta}(\hat{\bar{\theta}})$. (See Figure 2-9 below.)

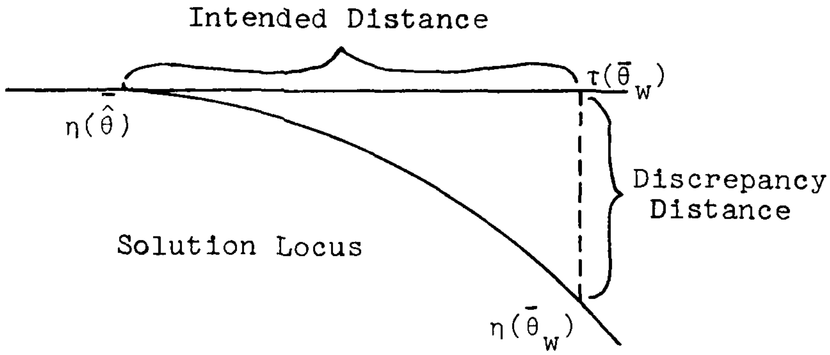


Figure (2-9). Illustration of Discripancy Distance Versus Intended Distance

In our case, if the degree of non linearity, $\hat{N}_{\theta}$, of the model, equation (2-3), satisfies inequality (2-39), then the linear theory results, with an appropriate correction factor, can be applied to find the confidence region for the non linear performance function for given values of $\bar{\theta}$ using

$$SS(\bar{\theta})-SS(\hat{\theta}) = ps^2F_{\alpha,p,N-p}(\text{correction factor}) \qquad (2\text{-}40)$$

Conversely, it will also be possible to find the confidence limits on the parameters, $\bar{\theta} = (a,b)$, by solving equation (2-40) when the confidence region for the function is specified. Going one step further, it will also be possible then to compute the confidence limits for the slope of the curve, where

$$\text{slope} = abt^{-(b+1)} \qquad (2\text{-}41)$$

at any trial, t_i, by conducting a search over the periphery of the joint confidence region of the parameters.

The correction factor discussed by Beale [4], in equation (2-40) when p=2 and s^2 is replaced by $SS(\bar{\theta})/(N-p)$ is,

$$[1 + \frac{N(p+2)}{(N-p)p} \hat{N}_\theta]$$

and includes the measure of non-linearity, $\hat{N}_\theta$, which allows for the effect on non linearity on the usual linear theory results.

As the difference between $\bar{n}(\bar{\theta}_w)-\bar{\tau}(\bar{\theta}_w)$, for a non-linear function (see Figure 2-9) increases within the region where the inequality sign in equation (2-39) does not change, the allowable size for the variance decreases. Therefore a

restriction using linear theory approximation is that the variance might have to be unreasonably small for the approximations to hold.

A study was conducted to determine the maximum size variance allowable for varying degrees of non-linearity of the performance curve function for a particular α level of significance. As the value of the coefficient or the exponent in equation (1-2) increases, the maximum allowable variance decreases. The results are contained in tables (3-18) and (3-19).

Procedure

To assist in following the non-linear procedure used to test for learning, an overview of the steps required is presented, followed by a detailed discussion in Appendix A.

Step 1: Estimate the parameters of the performance function and the variance using non-linear estimation techniques.

Step 2: Find the degree of non-linearity, $\hat{N}_\theta$, for the estimated function using Beale's measure of non-linearity.

i. If $\hat{N}_\theta \leq .01/F_{\alpha,p,N-p}$ proceed to step 3

ii. If $\hat{N}_\theta > .01/F_{\alpha,p,N-p}$ stop, following procedure not valid

Step 3: Determine the confidence limits for each Parameter of the performance function by satisfying equation (2-40) using a direct technique.

Step 4: Find the confidence limits for the slope at any particular trial using a direct search technique over the periphery of the joint parameter confidence region. The maximum and minimum values for the slope will be the upper and lower confidence limits respectively.

Step 5: Examine the confidence limits for the slope:

(i) If the confidence interval contains zero, do not reject the hypothesis that learning did not occur during this trial.

(ii) If the confidence interval does not contain zero, then it can be concluded that learning is taking place during this trial.

Steps 1-4 are explained in greater detail in Appendix C. A computer program is located in Appendix D. which estimates the parameters and the variance, tests the degree of nonlinearity and computes the confidence limits on the slope at any particular trial.

CHAPTER III

EVALUATION OF PROCEDURES

Since the accuracy of the test on the slope will depend on the quality of the estimate of σ_ε^2, a computer simulation was conducted to compare the estimators of σ_ε^2 presented in Chapter II. By varying the parameters for any given sample size of observations, we can simulate different situations that could occur in an actual test.

Evaluation of the Bias in Estimating σ_ε^2

The expected bias value was computed first for each estimator for specified parameter values and sample size. As the exponent value, b, or the coefficient, a, increased, there was a corresponding increase in the bias value associated with each estimator (see Figures 3-1 through 3-6). This resulting increase in bias as the parameter values are increased, corresponds to a more severe lack of fit of the linear model to the actual process. The best estimator using minimum expected bias as the selection criteria, is equation (2-8). The expected bias of this estimator for given parameter values and sample size was approximately 40 percent less than the expected bias obtained using the next best estimator, equation (2-9). The expected bias associated with equation (2-7) was 4 times larger than the bias factor using equation (2-8).

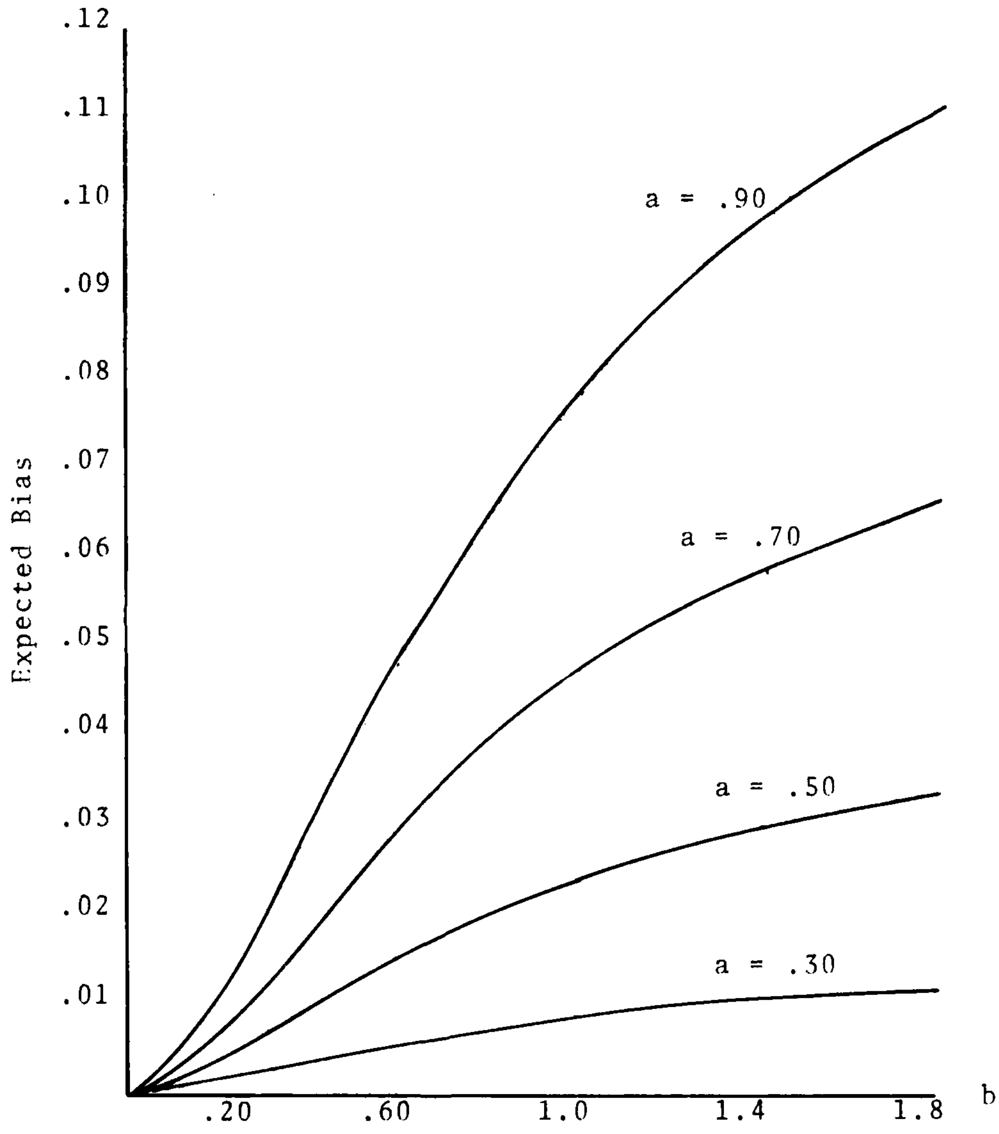


Figure 3-1. Expected Bias Using Equation (2-7) to Estimate σ_ε^2, N = 6.

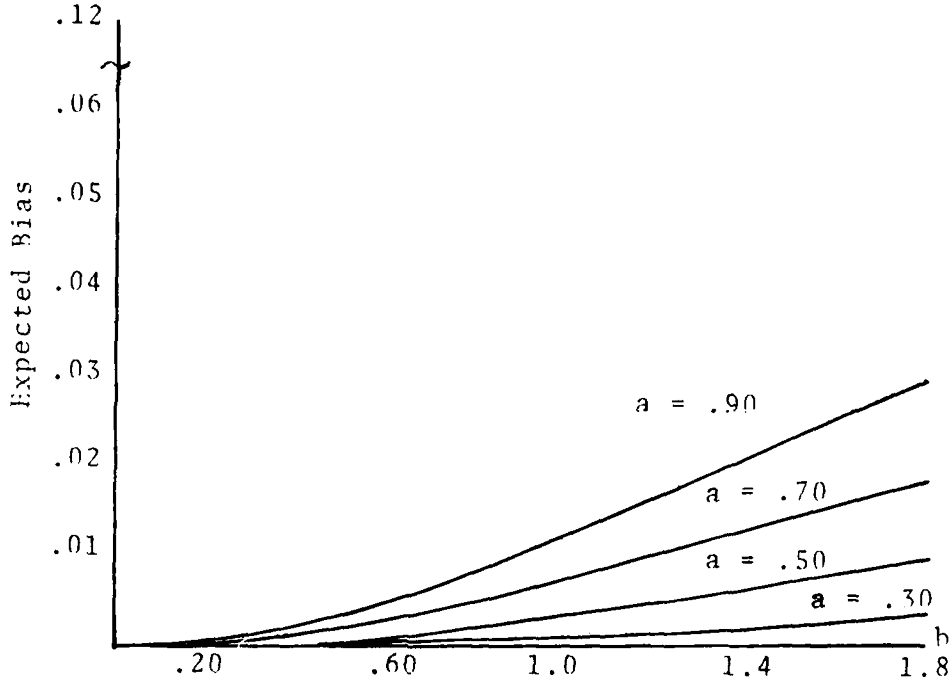


Figure 3-2. Expected Bias Using Equation (2-8) to Estimate σ_ε^2, N = 6.

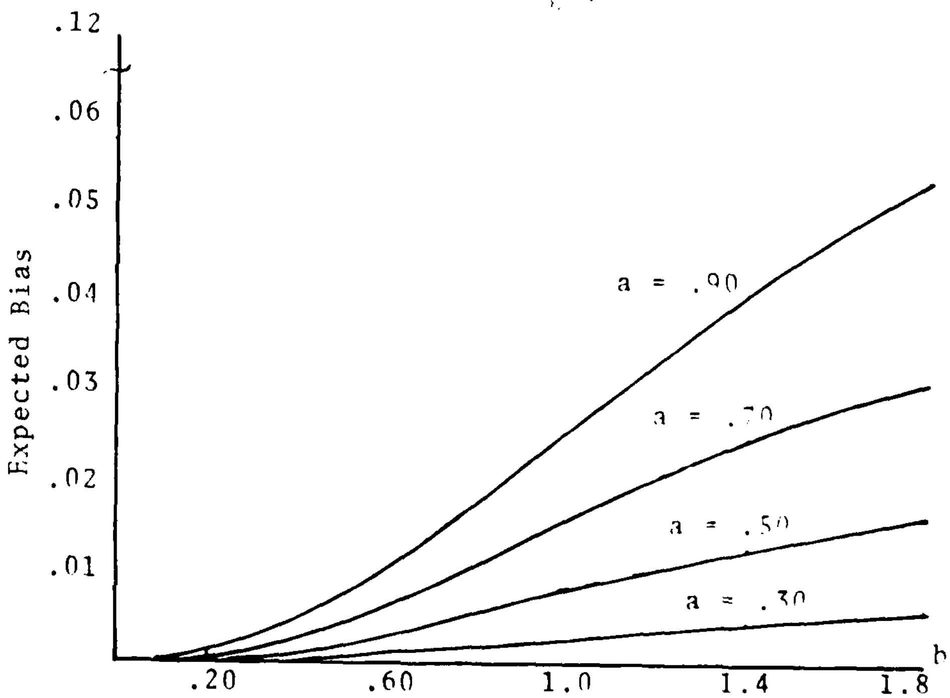


Figure 3-3. Expected Bias Using Equation (2-9) to Estimate σ_ε^2, N = 6.

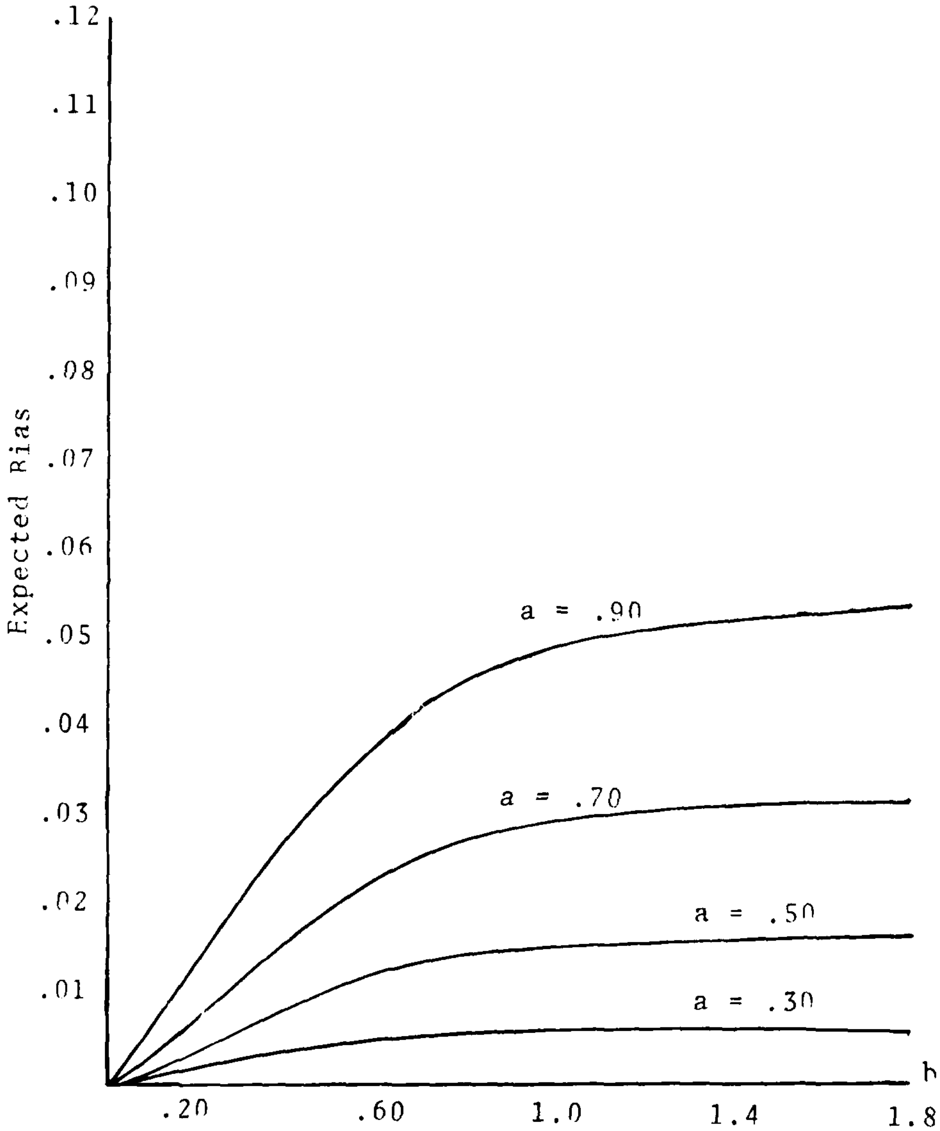


Figure 3-4. Expected Bias Using Equation (2-7) to Estimate σ_e^2, N = 15

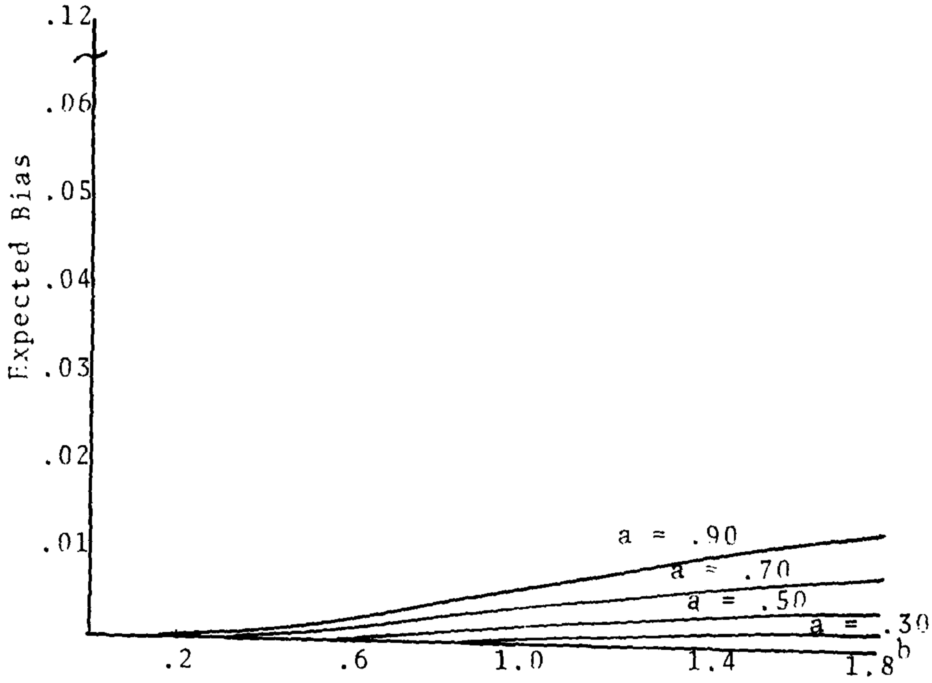


Figure 3-5. Expected Bias Using Equation (2-8) to Estimate $\sigma_\varepsilon{}^2$, N = 15

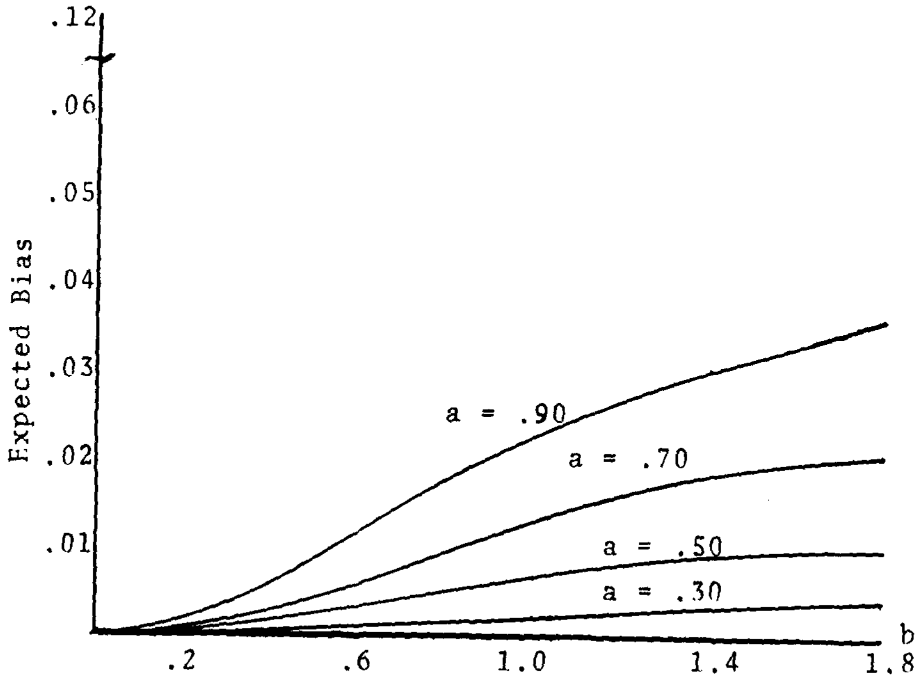


Figure 3-6. Expected Bias Using Equation (2-9) to Estimate $\sigma_\varepsilon{}^2$, N = 15

A simulation study was conducted to evaluate the variance of each estimator for all combinations of sample size N = 6, 15; standard deviation σ_ε = .03, .05, .07, .09; and parameter values a = .1, .3, .5 and b = 0, .4, .8, 1.2. To reduce sampling variability in the evaluation, the same stream of normal randomly generated observations was used for each estimator. One run consisted of 1000 experiments for the specified sample size, parameter values and variance. The results obtained are in terms of percent of the extimates within the specified tolerance limits ($\sigma_\varepsilon \pm \delta$), when using a particular estimator.

Designating any estimate as good or bad depending on whether it falls within or outside the tolerance limits respectively, then the generating process of a particular estimator will follow a binomial distribution. Using the worst case for estimating the variance of the process, (i.e. p = .5), we can approximate the variance as .25/N. To obtain a 95% confidence interval for the percent of good extimates, p, with limits ($\hat{p}$ ± .03), an appropriate sample size would be calculated as follows:

$$\hat{p} + .03 = \hat{p} + 2\sqrt{\hat{p}\hat{q}/N}$$

Letting p = .5

$$N = \frac{4pq}{(.03)^2}$$

$$N = 1100$$

The number of experiments for each combination was set at 1000. The actual variance for p may be less for a given condition since p = .5 will give the upper value of the variance for this process.

Since the number of simulation runs is sufficiently large, N = 1000, a normal approximation of the binomial variables can be used to construct a significance test [11]. To determine if the percent of good estimates using one estimator differs from the percent of good estimates using another estimator conduct the following test:

$$H_0: \ p_1 = p_2$$

$$H_1: \ p_1 \neq p_2$$

Compute

$$Z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{N} + \frac{\hat{p}_2(1-\hat{p}_2)}{N}}} \qquad (3\text{-}1)$$

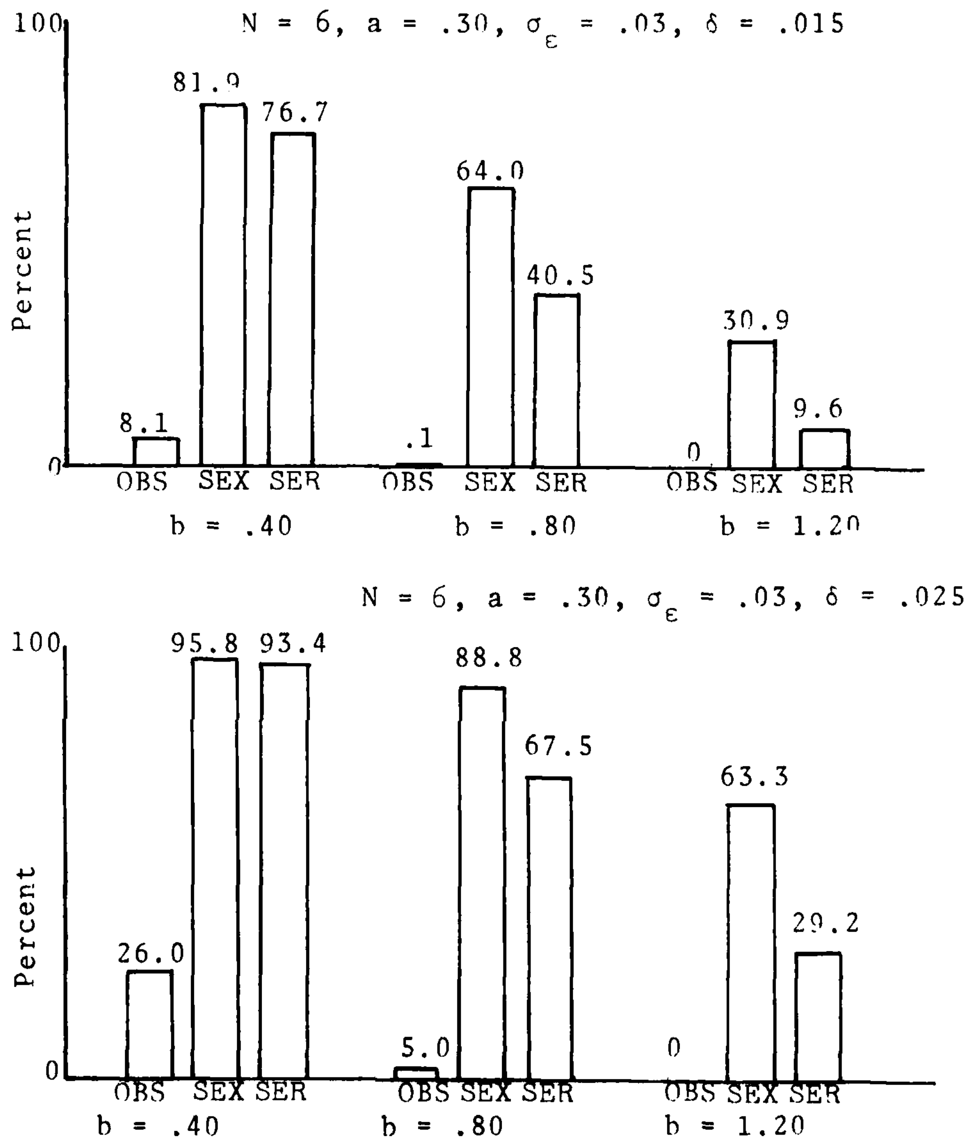
If $Z_0 > Z_{\alpha/2}$ reject H_0 and assume there is a significant difference in the p values for the two estimators.

The actual variance of the difference between the two success ratios will be less than $\frac{p_1(1-p_1)+p_2(1-p_2)}{N}$ due to correlation between the estimates from the two estimators. Although the assumption of independence between the variances of the success ratios, p_1 and p_2, reduces the power of the test, we can gain some idea of the significance of the difference between the two estimators using equation (3-1).

The results and analysis of this study on the estimators of σ_ε follow. When the true variance is small ($\sigma_\varepsilon^2 \leq .0036$), the best estimator is equation (2-8), regardless of the sample size or parameter values of the curve. As the variance increases however, equation (2-9) appears to perform better under certain conditions than does equation (2-8). See Figures (3-7) through (3-14). To examine this situation further, the distribution of the extimates of the variance about their respective expected values was analyzed for different conditions. The effect of sample size appears to be the same on each estimator, (i.e. larger N results in smaller estimator variance) and is not significant in determining the best estimator.

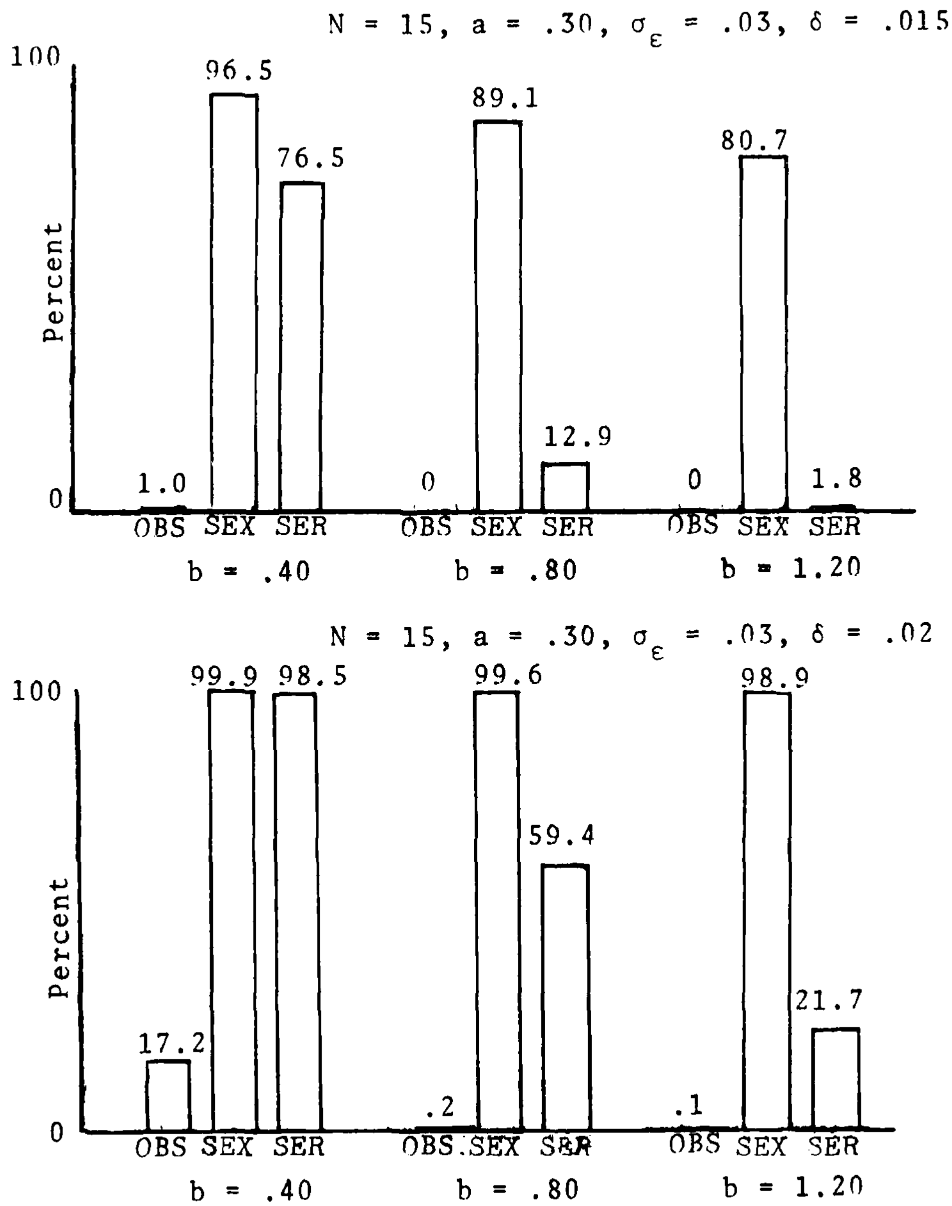
A general description of the distribution of the estimates using equation (2-8) and equation (2-9) for a 2^3 design in terms of σ_ε and parameters "a" and "b" is depicted in Figures 3-15 through 3-22.

The results noted were:



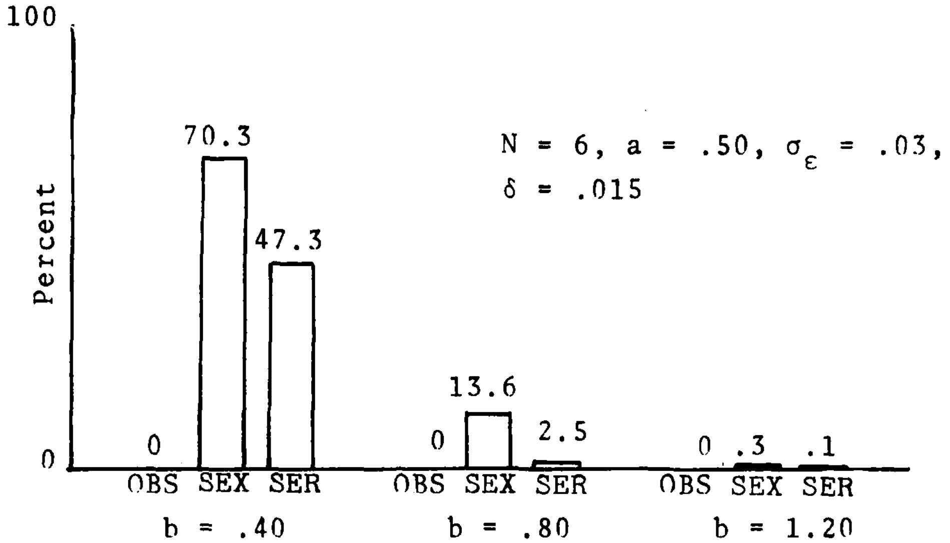
OBS represents estimates obtained using equation (2-7)
SEX represents estimates obtained using equation (2-8)
SER represents estimates obtained using equation (2-9)

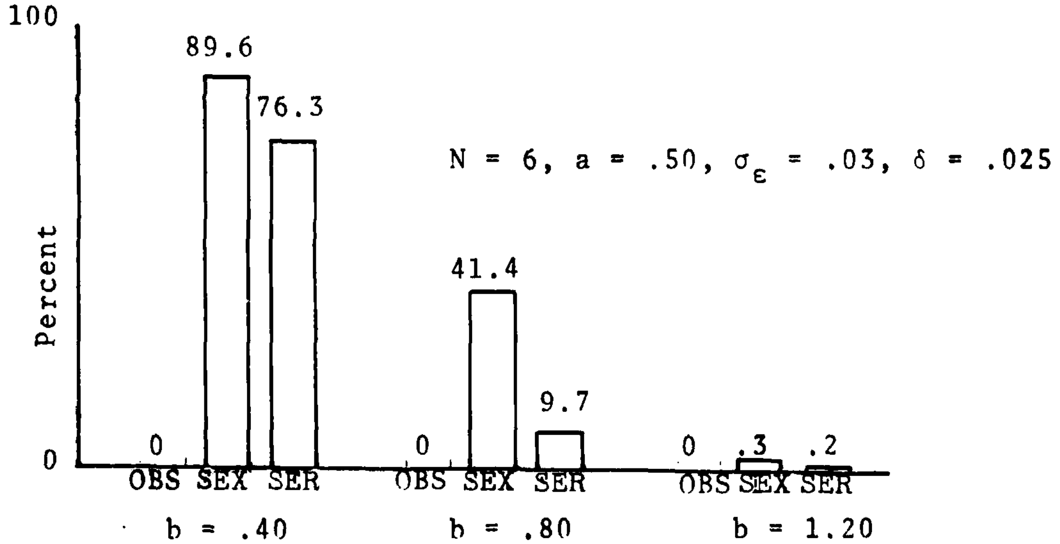
Figure 3-7. Percent of Estimates Within the Interval ($\sigma_\varepsilon \pm \delta$) when N, σ_ε, a, b, and δ are Specified



OBS represents estimates obtained using equation (2-7)
SEX represents estimates obtained using equation (2-8)
SER represents estimates obtained using equation (2-9)

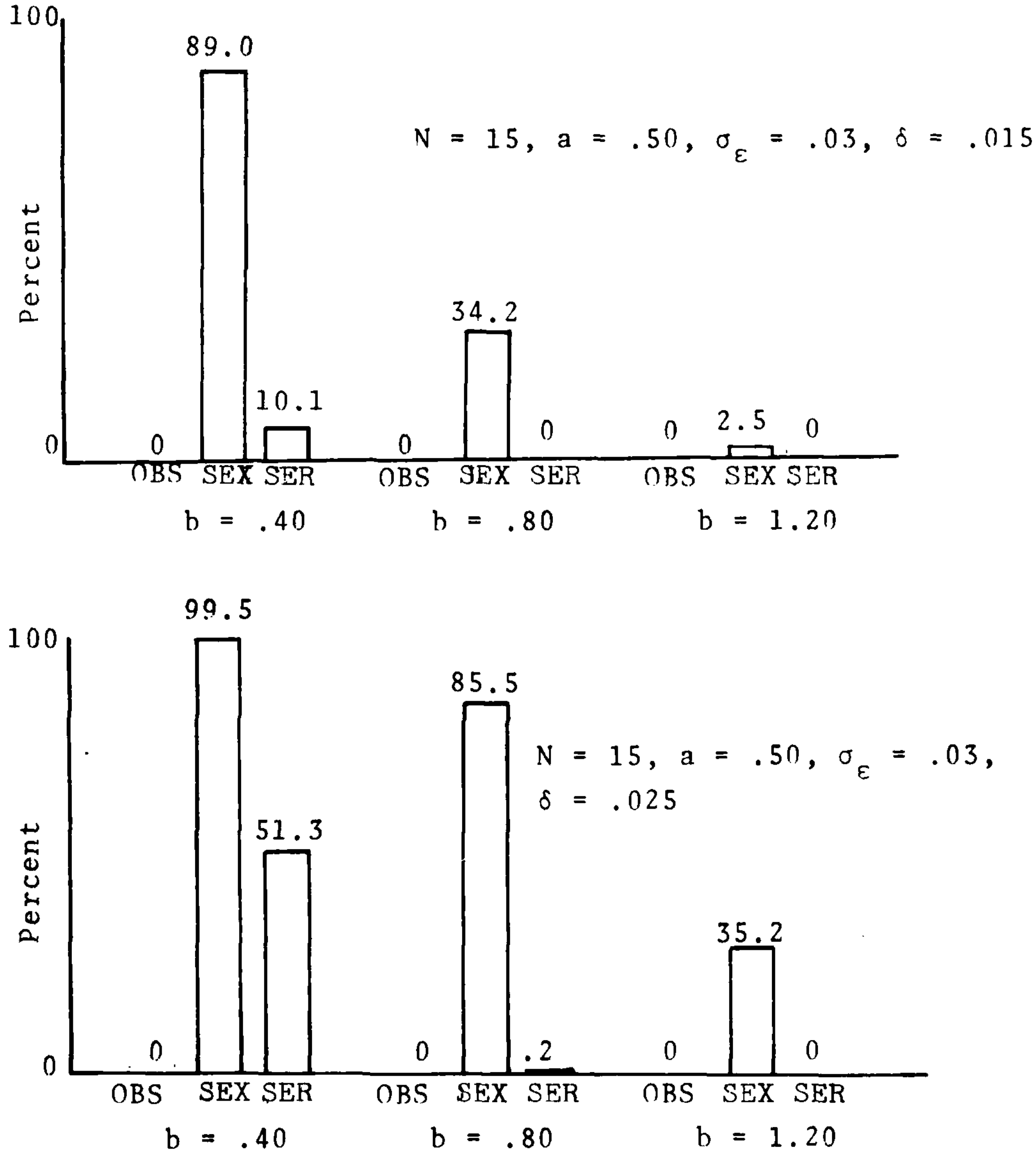
Figure 3-8. Percent of Estimates Within the Interval ($\sigma_\varepsilon \pm \delta$) when N, σ_ε, a, b, and δ are Specified





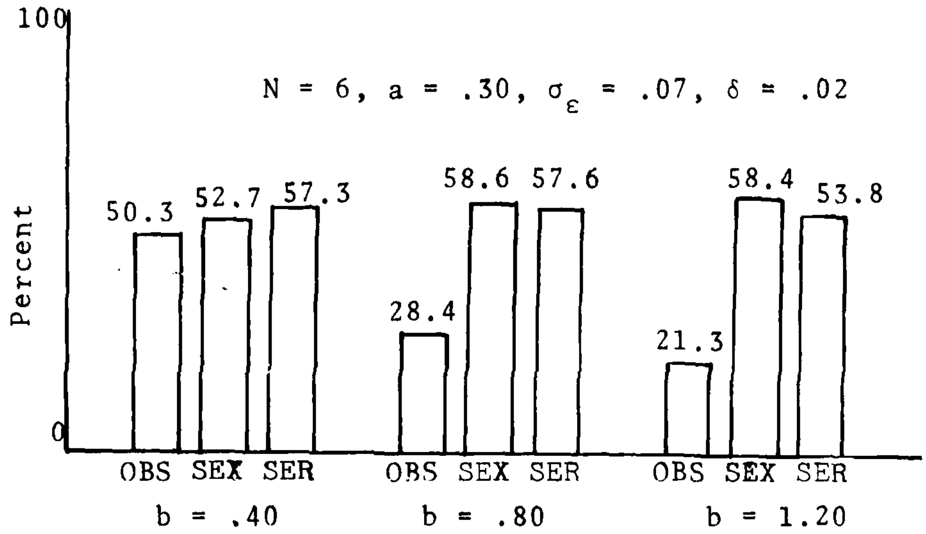
OBS represents estimates obtained using equation (2-7)
SEX represents estimates obtained using equation (2-8)
SER represents estimates obtained using equation (2-9)

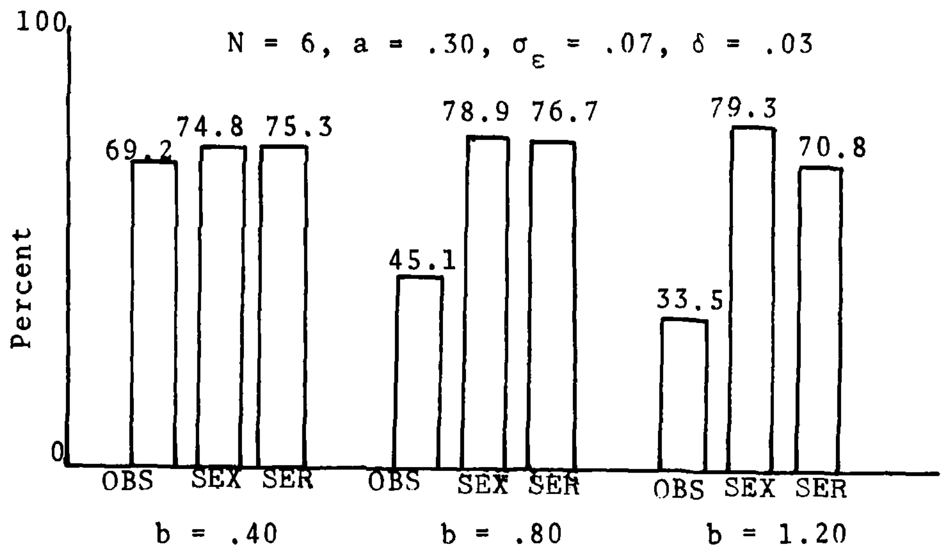
Figure 3-9. Percent of Estimates Within the Interval ($\sigma_\varepsilon \pm \delta$) when N, σ_ε, a, b, and δ are Specified



OBS represents estimates obtained using equation (2-7)
SEX represents estimates obtained using equation (2-8)
SER represents estimates obtained using equation (2-9)

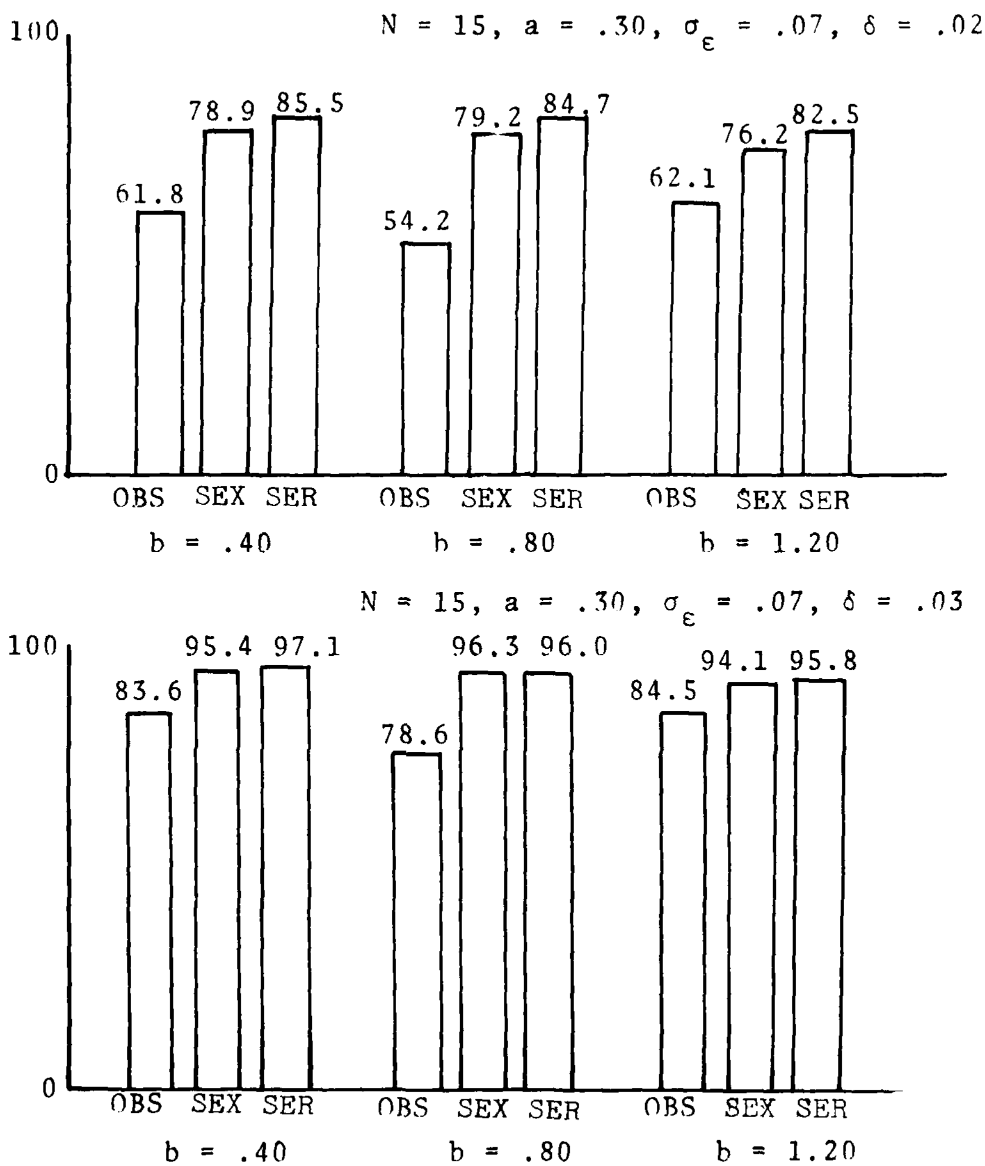
Figure 3-10. Percent of Estimates Within the Interval ($\sigma_\varepsilon \pm \delta$) when N, σ_ε, a, b, and δ are Specified





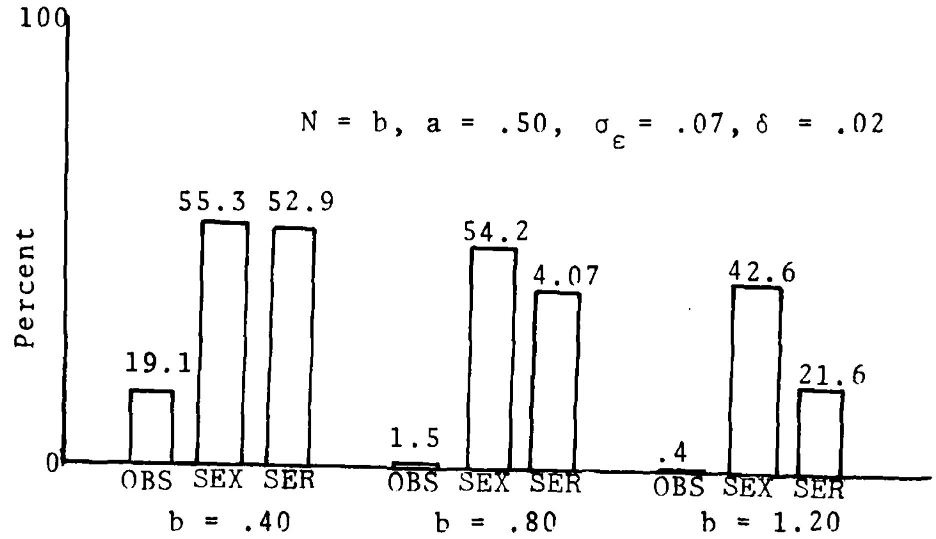
OBS represents estimates obtained using equations (2-7)
SEX represents estimates obtained using equations (2-8)
SER represents estimates obtained using equations (2-9)

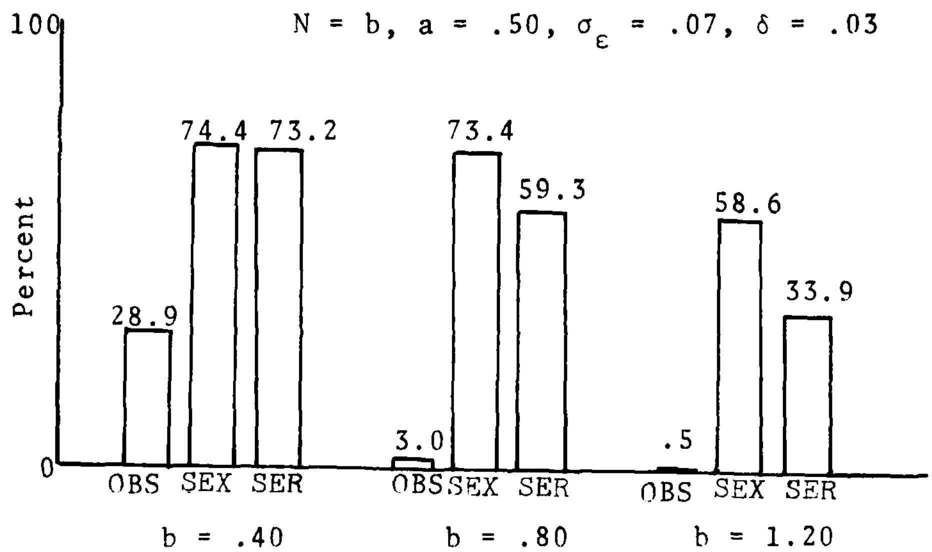
Figure 3-11. Percent of Est-mates Within the Interval ($\sigma_\varepsilon \pm \delta$) when N, σ_ε, a, b, and δ are Specified



OBS represents estimates obtained using equation (2-7)
SEX represents estimates obtained using equation (2-8)
SER represents estimates obtained using equation (2-9)

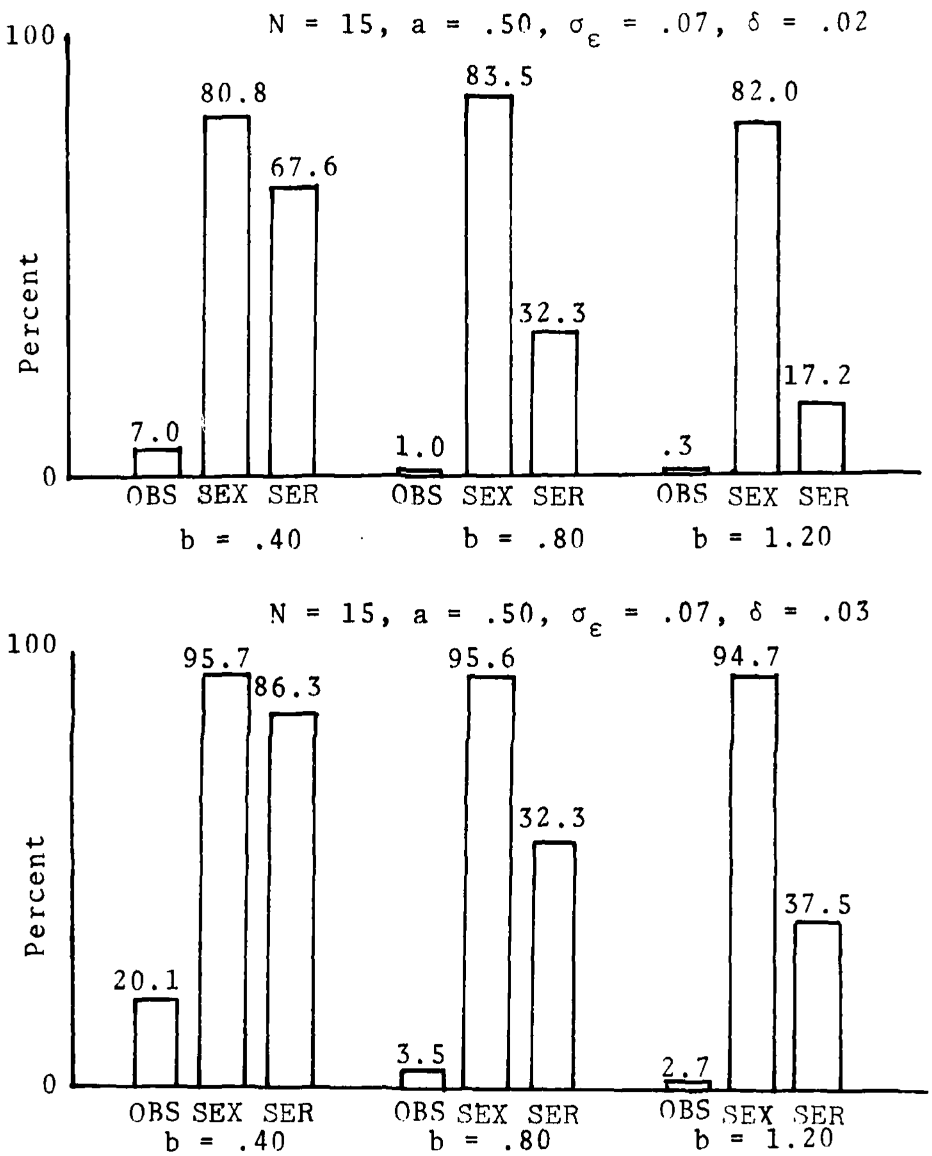
Figure 3-12. Percent of Estimates Within the Interval ($\sigma_\varepsilon \pm \delta$) when N, σ_ε, a, b, and δ are Specified





OBS represents estimates obtained using equations (2-7)
SEX represents estimates obtained using equations (2-8)
SER represents estimates obtained uding equations (2-9)

Figure 3-13. Percent of Estimates Within the Interval ($\sigma_\varepsilon \pm \delta$) When N, σ_ε, a, b, and δ are Specified



OBS represents estimates obtained using equation (2-7)
SEX represents estimates obtained using equation (2-8)
SER represents estimates obtained using equation (2-9)

Figure 3-14. Percent of Estimates Within the Interval ($\sigma_\varepsilon \pm \delta$) when N, σ_ε, a, b, and δ are Specified

(i) As the variance of the process increased for large values of parameter b, the dispersion of the estimates using either estimator appears to be a heavy tailed normal distribution. This should be expected since larger estimates of the variance are possible which will weigh the expected value of the estimates to the right. It is also worthwhile to point out that when parameter "a" is small in this situation, the majority of the estimates obtained were less than the true variance. Thus it appears that near the asymptote, when the variance is large, the linear model approximation using the difference method results in a "better fit" than the true model in terms of minimizing the sum of the squared errors.

(ii) When the variance of the process was large and both parameters small, the estimates using equation (2-8) appeared to be more dispersed about their expected value than the estimates using equation (2-9) about their expected value. When the expected value of the estimates is near the true value of the variance, a tighter distribution of the estimates about their expected value can result in more estimates falling within the tolerance limits (see Figure 3-17). For the two situations discussed in (i) and (ii) above, the estimator with the largest expected bias generates more estimates within the tolerance limits specified for this study. It is easy to see in Figures 3-17 and 3-18 that if the tolerance limits were reduced it becomes more difficult to differentiate between which estimator is better.

For all other combinations of parameters and variance size examined, the best estimator was the one with the minimum expected bias. The rule for selecting which estimator of σ_ε is better when the parameters and the true value of σ_ε are known is:

If:

(i) $\sigma_\varepsilon \leq .06$ Use equation (2-8) to estimate σ_ε

(ii) $\sigma_\varepsilon > .06$, $a \geq .5$ Use equation (2-8) to estimate σ_ε

(iii) $\sigma_\varepsilon > .06$, $a < .5$ Use equation (2-9) to estimate σ_ε

Since the true value of the parameters and the actual variance will not be known, we must base our rule on estimates for and parameter "a". A general rule then for selecting the best estimator when σ_ε and parameter "a" must be estimated is:

Compute an estimate of σ_ε using equation (2-8)

$$\hat{\sigma}_\varepsilon = \sqrt{\frac{(N-1)\sum_{i=1}^{N-1}(x_i-\bar{x})^2}{2N(N-1)}}$$

If

(i) $\hat{\sigma}_\varepsilon \leq .06$ Use $\hat{\sigma}_\varepsilon$ as computed

(ii) $\hat{\sigma}_\varepsilon > .06$, $\hat{a} \geq .5$ Use $\hat{\sigma}_\varepsilon$ as computed

(iii) $\hat{\sigma}_\varepsilon > .06$, $\hat{a} < .5$ Use equation (2-9) to compute an estimate σ_ε

[illegible]ffectiveness of this rule, when choosing an estimator [illegible] will be evaluated in conjunction with determining

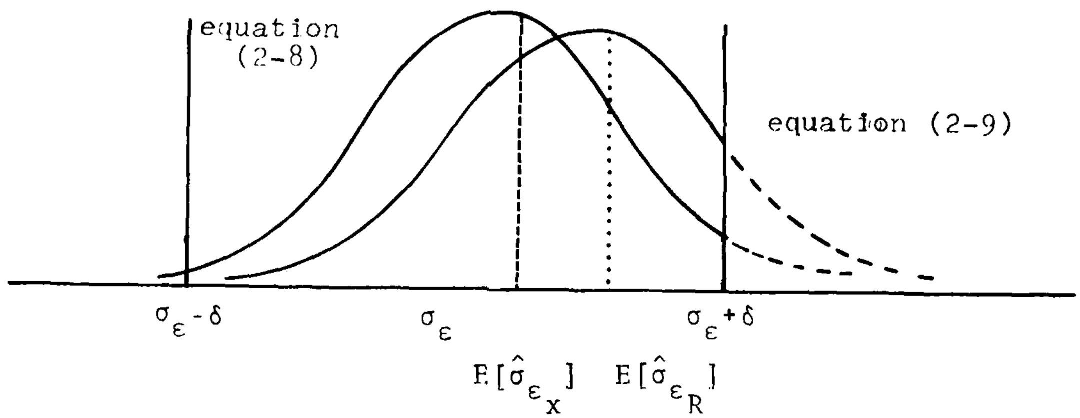


Figure 3-15. General Description of the Distribution of the Estimates of σ_ε Using Equations (2-8) and (2-9) when σ_ε is Small, a is Small, b is Small

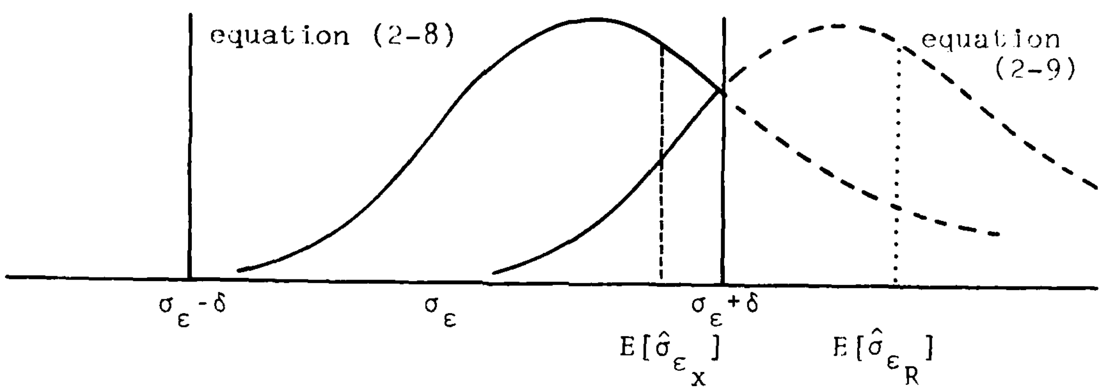


Figure 3-16. General Description of the Distribution of the Estimates of σ_ε Using Equations (2-8) and (2-9) when σ_ε is Small, a is Small, b is Large

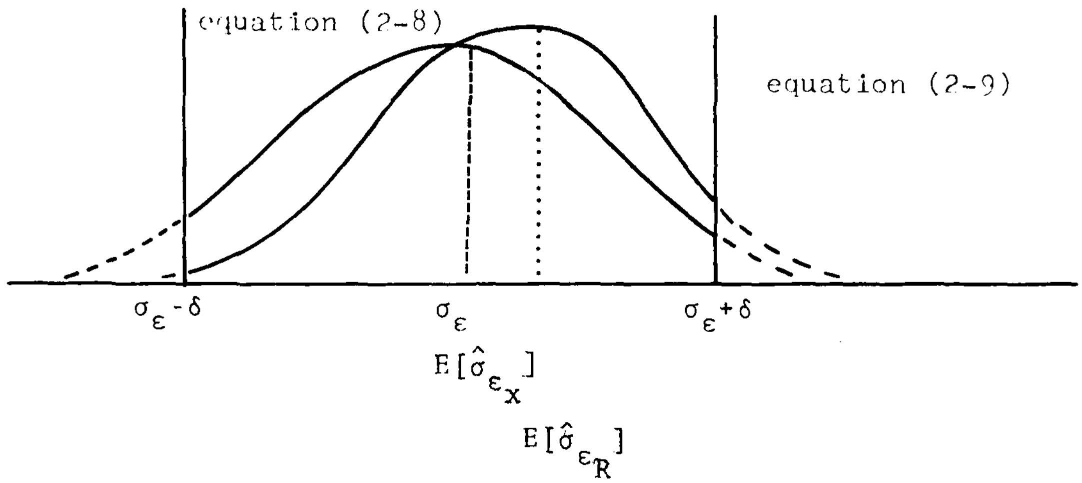


Figure 3-17. General Description of the Distribution of the Estimates of σ_ε Using Equations (2-8) and (2-9) when σ_ε is Large, a is Small, b is Small

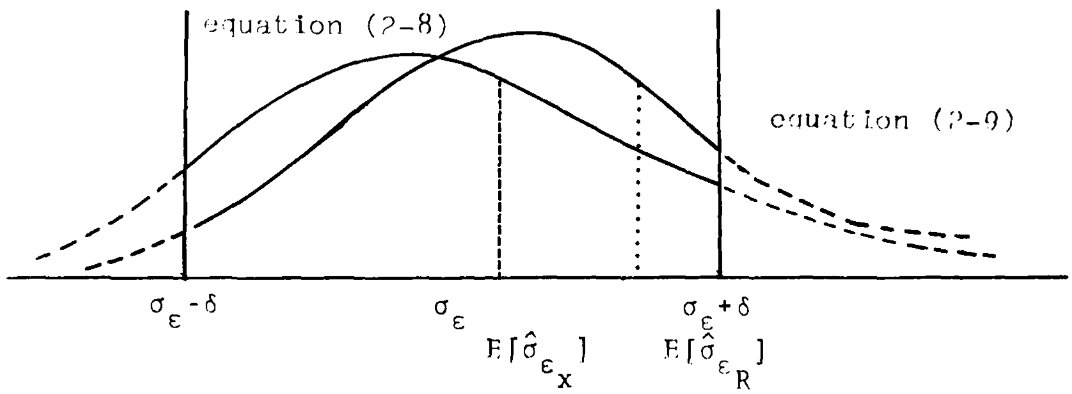


Figure 3-18. General Description of the Distribution of the Estimates of σ_ε Using Equations (2-8) and (2-9) when σ_ε is Large, a is Small, b is Large

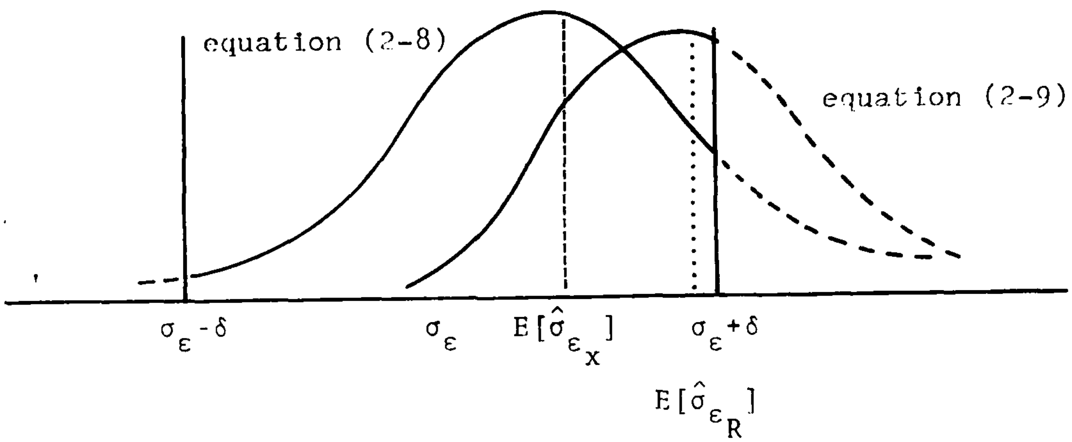


Figure 3-19. *General Description of the Distribution of the* Estimates of σ_ε Using Equations (2-8) and (2-9) when σ_ε is Small, a is Large, b is Small

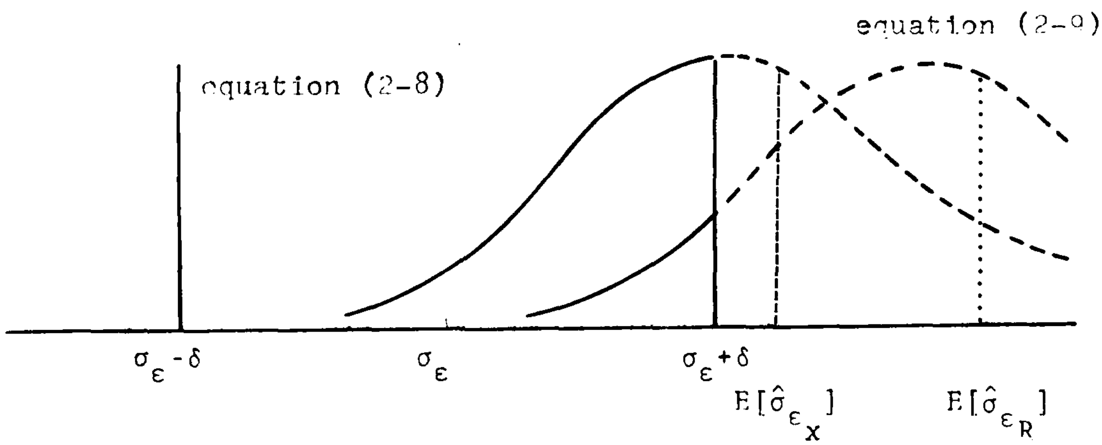


Figure 3-20. General Description of the Distribution of the Estimates of σ_ε Using Equations (2-8) and (2-9) when σ_ε is Small, a is Large, b is Large

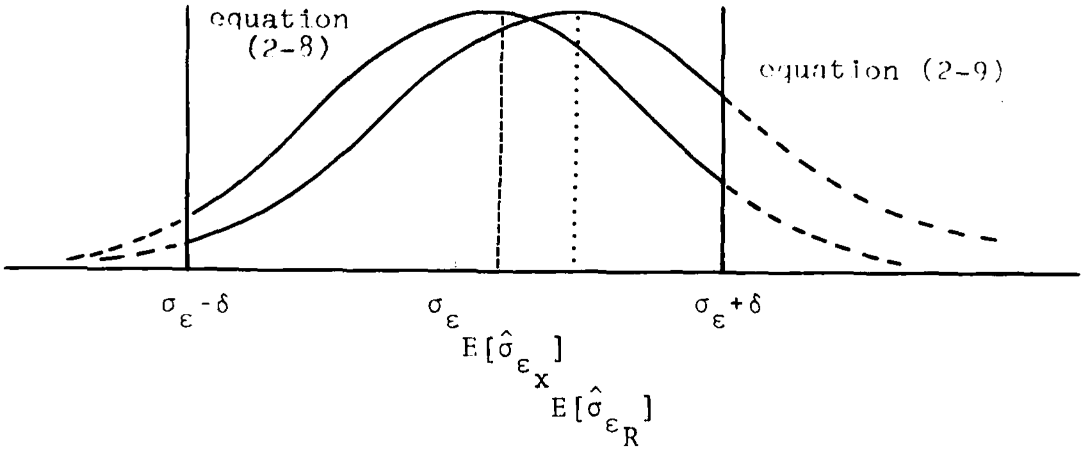


Figure 3-21. General Description of the Distribution of the Estimates of σ_ε Using Equations (2-8) and (2-9) when σ_ε is Large, a is Large, b is Small

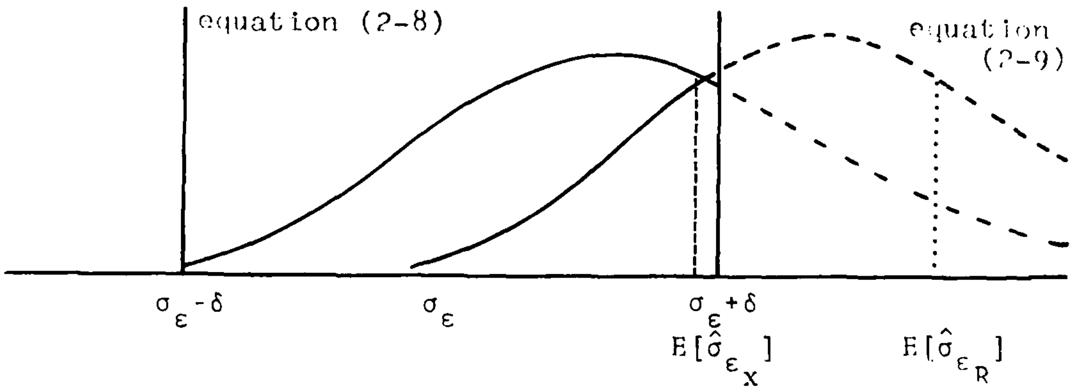


Figure 3-22. General Description of the Distribution of the Estimates of σ_ε Using Equations (2-8) and (2-9) when σ_ε is Large, a is Large, b is Large

the best test procedure later in this chapter.

Comparison of the Linear Test Methods

Test Learning

The LLSR estimate of the average rate of learning as defined by equation (2-23) was analyzed in terms of the percent of the true average rate of learning for various combinations of sample size and parameter values. If the size of the sample is increased, the amount of bias in the estimate of the average rate of learning, equation (2-20) using the LSSR method should also increase in a negative direction. Stated another way, as the sample size increases, the expected value of the LLSR estimate of average rate of learning decreases. This should be expected since the observations over the latter trials will be closer to the asymptotic value which will have a negative effect on the slope of the LLSR model.

As can be seen in Figures 3-23 and 3-24 the effect of N on the expected value of the average slope estimate using the LLSR method is more pronounced than its effect on the ACD slope estimate, which is unbiased.

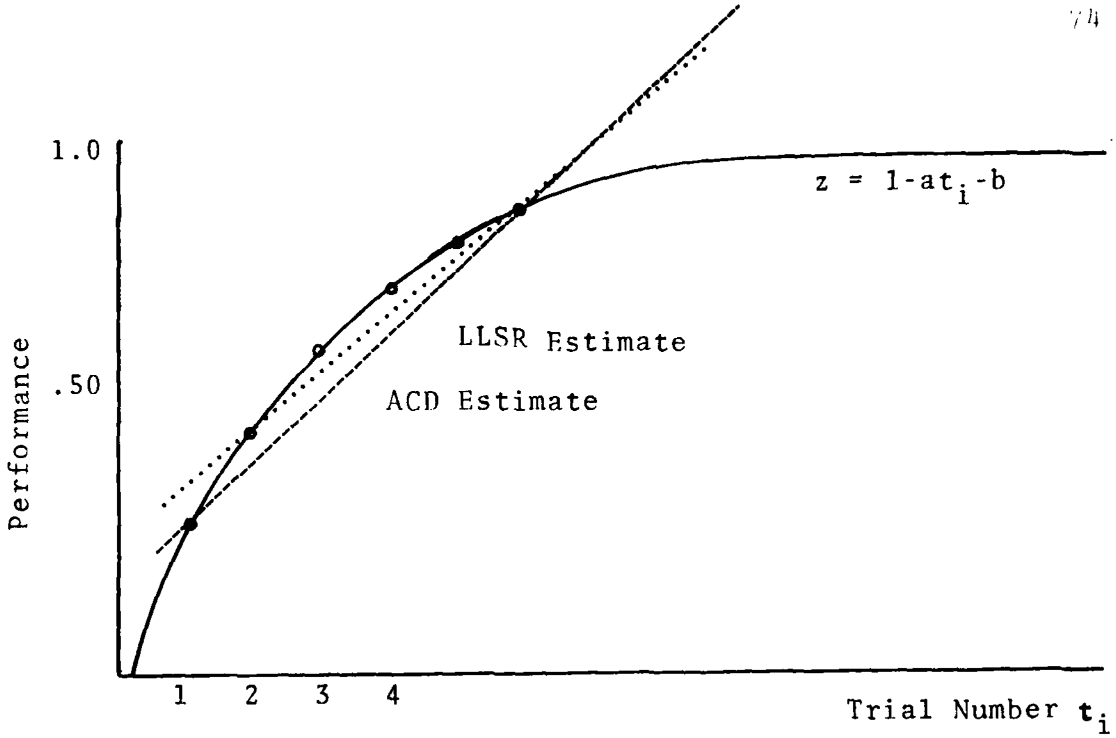


Figure 3-23. LLSR Estimate of Slope in Relation to Average Rate of Learning, Also the ACD Estimate, when N is Small

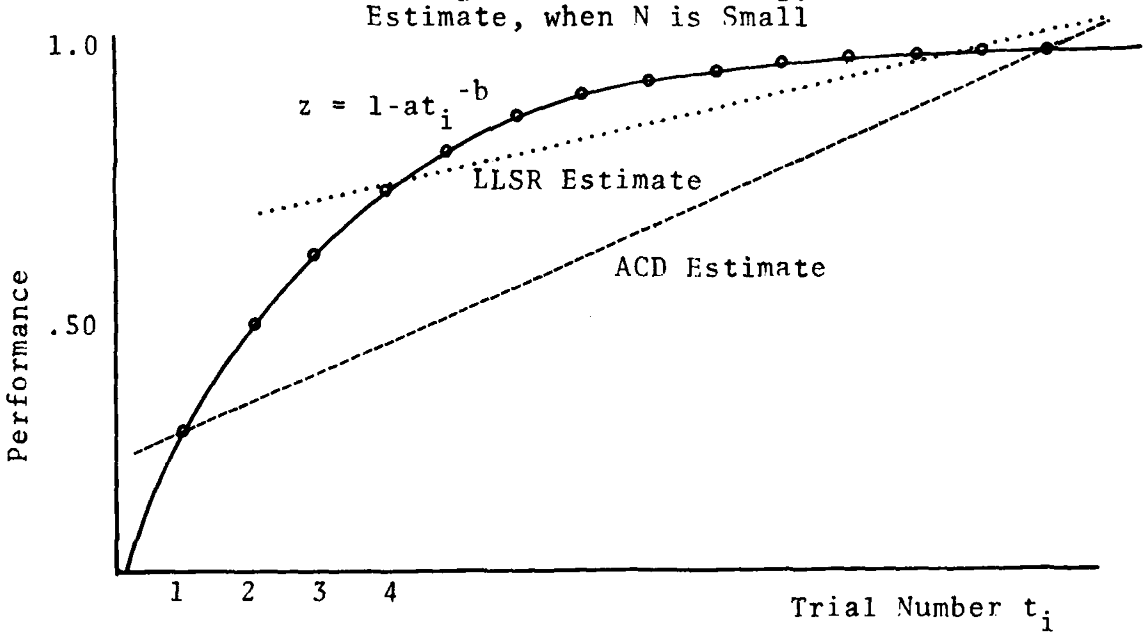


Figure 3-24. LLSR Estimate of Slope in Relation to Average Rate of Learning, Also the ACD Estimate, when N is Large

In Figure 3-24, if the value of parameter b approaches zero, the ratio of the LLSR estimate to the average slope will approach 1. When the value of b is zero, this is equivalent to the expected values of the observations being at the asymptote of the performance curve. We showed in Chapter II that under this condition the LLSR estimate will be unbiased. Therefore when the value of b is small, the expected value of the LLSR slope estimate will be very close to the actual average rate of learning; and as the value of b increases, the expected difference between the LLSR estimate and the true average slope will increase.

The effect of parameter a is not as intuitively obvious. Since parameter "a" can be factored out of both equation (2-23) and equation (2-20), we see that the expected value of the LLSR estimate and the true average slope change by an equal multiple factor. Therefore, although the LLSR estimate and the actual average slope values change, the ratio of the two values is not affected. This ratio is graphed in Figure 3-25 as a function of the exponent value, b, and sample size, N.

To evaluate the significance of the bias in the LLSR estimate when testing for learning, it will be necessary to examine the effect on the computed test statistic. Recall that the test statistic for the LLSR method is

$$t_o = \frac{d_{LLSR} - 0}{\sqrt{\dfrac{12\sigma_\varepsilon^2}{N(N^2+1)}}}$$

and for the ACD method is

$$t_o = \frac{d_{ACD} - 0}{\sqrt{\dfrac{2\sigma_\varepsilon^2}{(N-1)^2}}}$$

Using the same estimate of the process variance when computing the slope variance for the LLSR and the ACD methods respectively, a comparison of the expected values of the test statistics was made. The ratio of the test statistics is graphed in Figure 3-26 as a function of the sample size, N. and the exponential value b. It should not be surprising to find that when the value of b is small for a particular sample size, the ratio of the LLSR test statistic to the ACD test statistic is greater than 1. This is because the ratio of the estimates for the average rate of learning using the LLSR and ACD methods is greater than the ratio of their respective variances. As the value of the exponent, b, increases for given N, the ratio of the average slope estimates becomes smaller and the ratio of the test statistics decrease.

An increase in the sample size for any particular

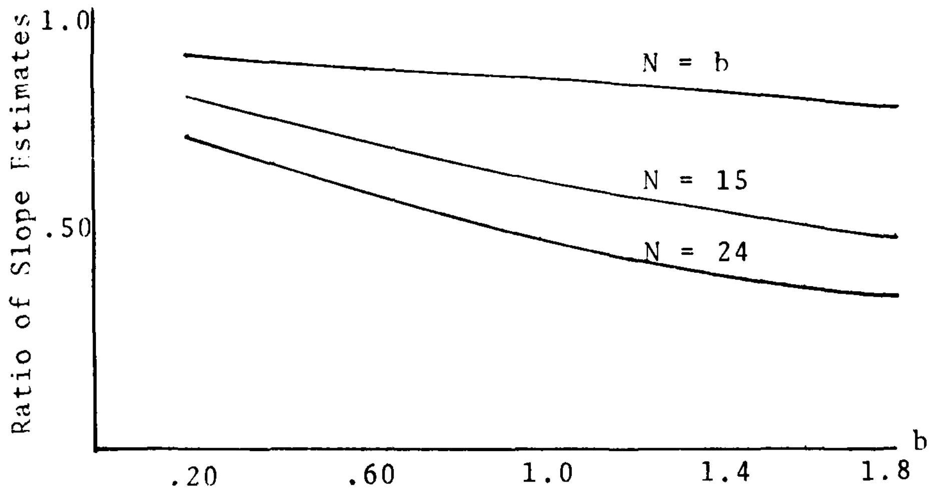


Figure 3-25. Ratio of the Estimates of the Average Slope Using the LLSR Method Versus ACD Method

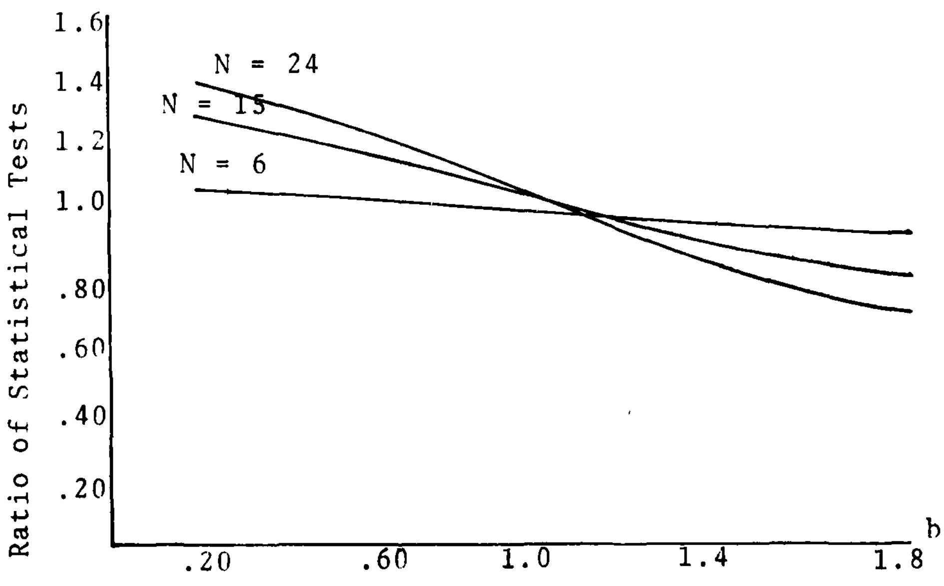


Figure 3-26. Ratio of Expected Test Statistics Using LLSR Test Procedure Versus the ACD Test Procedure

value of b will also affect the ratio of the test statistics. This is due to the fact that both the LLSR slop estimate and the corresponding standard deviation of this slope estimate decrease at a more rapid rate than the ACD slope estimate and its corresponding standard deviation for increases in N (see Figure(3-26). If the ratio of the slope estimates, LLSR vs ACD, decreases more rapidly than the ratio of the corresponding standard deviations of these estimates, then when N and b assume certain values, the ACD test procedure will become more powerful. Thus if we knew the parameter value b, we could select the more powerful test procedure by examining the sample size N. Since b will usually be unknown and difficult to estimate without using a computer search technique, we need a general rule based only on the sample size.

Recall from Chapter I, that our purpose for developing the test procedures was to test if learning is occurring during an evaluation of a particular system. Since the people operating the system have undergone extensive training prior to the evaluation, it seems unlikely that the rate of learning that could occur during the trials of this evaluation would be very large. Therefore the performance curve measuring learning during an evaluation should rarely have a parameter value of $b \geq 1.0$. Based on this consideration, and referring back to Figure 3-26, we can expect the best linear test procedure for detecting learning to be the LLSR method since the expected value of the test statistic for $b \lesssim 1.0$ is greater than that for the ACD method when $N \leq 25$.

Evaluation of Linear Test Procedures

A simulation study was conducted to evaluate the findings in the previous section concerning the more powerful test procedure for all combinations of sample size, N = 6, N = 15; σ_ε = .03, .05, .07, .09, and parameter values a = .1, .3, .5 and b = 0, .4, .8, 1.2. The test procedure that results in the largest percent of significant test statistics for a given α level when learning is actually occurring will be selected as the better test. The study was also designed to evaluate if any one particular method of the three alternatives considered below was better for estimating σ_ε.

1. Using equation (2-8) under all conditions
2. Use equation (2-9) under all conditions
3. Apply the rule for choosing an estimator based on estimates of σ_ε and parameter "a" as discussed in the previous section.

The best method of the three alternatives will be selected based on the largest percent of runs that a particular alternative resulted in an estimate of σ_ε that was closest to the true value. One thousand simulation runs were used for each combination of σ_ε, sample size and parameter values to insure that the true percent of detection of learning would be within ±3% of the simulation results at least 95% of the time. To reduce the sampling variability in the evaluation,

the same stream of normal randomly generated observations was used in evaluating each test procedure for all combinations of $a, b, \sigma_{\varepsilon}, N$.

The results of the first part of the simulation study, the evaluation of the test procedures, are given in Tables 3-1 through 3-8. The labels SEX, SER, and RULE correspond to methods 1, 2, and 3, respectively, used to obtain an estimate of σ_{ε}. Each block in the table contains the percent of times the LLSR test detected learning, t_R; and the percent of times the ACD test detected learning, t_D, using a particular method for estimating σ_{ε}. The test statistics for each test procedure were compared at the $\alpha = .05$ level.

The results substantiate our earlier findings that the ratio of the LLSR test statistic to the ACD test statistic is a function of the sample size and parameter b. In Figure 3-26 we found that for values of $b \leq 1.0$ and sample size of $N = 6$ and $N = 15$, we should expect the LLSR test procedure to be more powerful than the ACD test. The results in Tables 3-1 through 3-8 support this. As the value of b increases from one the ACD test procedure is more powerful and the results also verify this. The type I error appears to be larger using the LLSR test procedure than the type I error using the ACD test procedure. This is expected since the LLSR test procedure gets more powerful in comparison to

the ACD test procedure as b decreases (see Figure 3-10). Also note however, that the type I error for the LLSR test procedure is greater than the specified α level but decreases as N increases. A possible explanation for this is that when parameters a and b are small, the estimates of σ_ε are approximately equally distributed about σ_ε as shown in Figures 3-15 and 3-17. If the estimate of σ_ε is smaller than σ_ε, which is the case over 50% of the time, we increase the probability of obtaining a type I error. As the sample size increases, the dispersion of the estimates of σ_ε about the true value of σ_ε becomes tighter and we have fewer estimates of σ_ε which are substantially less than the true value of σ_ε, thus decreasing the probability of a type I error. When the type one error was specified at α = .10 the power of both test procedures increased. See Table 3-9.

The second objective of the simulation was to determine if one particular method for estimating σ_ε was better under certain conditions than another. The results are shown in Tables 3-10 through 3-17. Using either equation (2-8) or equation (2-9) instead of applying the rule, yielded a larger percent of better estimates for almost all of the combinations examined. It appears then, that the estimates of σ_ε and parameter "a" are not accurate enough to use in applying our general rule.

AD-A086 176 GEORGIA INST OF TECH ATLANTA F/G 5/9
TESTING FOR LEARNING WITH SMALL DATA SETS.(U)
JUN 78 K A YEALY DAAG39-77-C-0041

UNCLASSIFIED NL

2 OF 2
ADA
086176

END
DATE
FILMED
8-80
DTIC

Table 3-1. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimate of σ_ε for Given Values of a, b, N, and σ_ε. α = .05

N = 6 σ_ε = .03

		b=0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.069 t_D=.042	t_R=.998 t_D=.995	t_R=1.000 t_D=1.000	t_R=.999 t_D=1.000
	SER	t_R=.060 t_D=.040	t_R=.994 t_D=.982	t_R=1.000 t_D=.999	t_R=.998 t_D=.996
	RULE	t_R=.069 t_D=.042	t_R=.997 t_D=.994	t_R=1.000 t_D=1.000	t_R=.999 t_D=.997
a=.3	SEX	t_R=.059 t_D=.047	t_R=.913 t_D=.859	t_R=.983 t_D=.974	t_R=.982 t_D=.985
	SER	t_R=.058 t_D=.039	t_R=.875 t_D=.813	t_R=.940 t_D=.929	t_R=.939 t_D=.918
	RULE	t_R=.059 t_D=.047	t_R=.914 t_D=.862	t_R=.973 t_D=.966	t_R=.959 t_D=.950
a=.1	SEX	t_R=.056 t_D=.043	t_R=.318 t_D=.272	t_R=.467 t_D=.415	t_R=.497 t_D=.452
	SER	t_R=.054 t_D=.038	t_R=.291 t_D=.256	t_R=.442 t_D=.395	t_R=.446 t_D=.409
	RULE	t_R=.056 t_D=.043	t_R=.318 t_D=.272	t_R=.467 t_D=.415	t_R=.497 t_D=.452

Table 3-2. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimate of σ_ε for Given Values of a, b, N, and σ_ε. α = .05

N = 6 σ_ε = .05

		b=0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.086 t_D=.067	t_R=.958 t_D=.930	t_R=.997 t_D=.991	t_R=.995 t_D=.997
	SER	t_R=.088 t_D=.070	t_R=.947 t_D=.914	t_R=.991 t_D=.980	t_R=.982 t_D=.983
	RULE	t_R=.088 t_D=.067	t_R=.958 t_D=.936	t_R=.995 t_D=.988	t_R=.990 t_D=.990
a=.3	SEX	t_R=.080 t_D=.074	t_R=.715 t_D=.632	t_R=.912 t_D=.871	t_R=.894 t_D=.904
	SER	t_R=.076 t_D=.073	t_R=.693 t_D=.605	t_R=.863 t_D=.835	t_R=.837 t_D=.826
	RULE	t_R=.081 t_D=.074	t_R=.719 t_D=.636	t_R=.881 t_D=.856	t_R=.845 t_D=.849
a=.1	SEX	t_R=.084 t_D=.061	t_R=.257 t_D=.218	t_R=.328 t_D=.285	t_R=.381 t_D=.367
	SER	t_R=.077 t_D=.060	t_R=.251 t_D=.214	t_R=.323 t_D=.286	t_R=.374 t_D=.355
	RULE	t_R=.084 t_D=.063	t_R=.259 t_D=.219	t_R=.328 t_D=.288	t_R=.389 t_D=.364

Table 3-3. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimate of σ_ε for Given Values of a, b, N, and σ_ε. $\alpha = .05$

$N = 6 \qquad \sigma_\varepsilon = .07$

		b=0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.084 t_D=.081	t_R=.823 t_D=.762	t_R=.951 t_D=.933	t_R=.968 t_D=.966
	SER	t_R=.085 t_D=.078	t_R=.797 t_D=.732	t_R=.910 t_D=.898	t_R=.933 t_D=.924
	RULE	t_R=.083 t_D=.081	t_R=.826 t_D=.764	t_R=.934 t_D=.918	t_R=.948 t_D=.943
a=.3	SEX	t_R=.077 t_D=.069	t_R=.506 t_D=.436	t_R=.731 t_D=.684	t_R=.773 t_D=.754
	SER	t_R=.078 t_D=.066	t_R=.499 t_D=.440	t_R=.697 t_D=.636	t_R=.705 t_D=.699
	RULE	t_R=.078 t_D=.069	t_R=.509 t_D=.451	t_R=.705 t_D=.649	t_R=.708 t_D=.704
a=.1	SEX	t_R=.092 t_D=.089	t_R=.197 t_D=.159	t_R=.246 t_D=.221	t_R=.249 t_D=.225
	SER	t_R=.092 t_D=.079	t_R=.186 t_D=.151	t_R=.234 t_D=.214	t_R=.238 t_D=.222
	RULE	t_R=.095 t_D=.089	t_R=.196 t_D=.160	t_R=.241 t_D=.160	t_R=.241 t_D=.219

Table 3-4. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimate of σ_ε for Given Values of a, b, N, and σ_ε. $\alpha = .05$

$N = 6 \qquad \sigma_\varepsilon = .09$

		b=0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.080 t_D=.074	t_R=.685 t_D=.631	t_R=.868 t_D=.843	t_R=.886 t_D=.874
	SER	t_R=.083 t_D=.071	t_R=.658 t_D=.594	t_R=.834 t_D=.799	t_R=.831 t_D=.831
	RULE	t_R=.077 t_D=.073	t_R=.680 t_D=.623	t_R=.862 t_D=.834	t_R=.860 t_D=.865
a=.3	SEX	t_R=.089 t_D=.082	t_R=.395 t_D=.359	t_R=.596 t_D=.545	t_R=.591 t_D=.582
	SER	t_R=.082 t_D=.078	t_R=.395 t_D=.351	t_R=.581 t_D=.523	t_R=.558 t_D=.548
	RULE	t_R=.092 t_D=.080	t_R=.402 t_D=.355	t_R=.582 t_D=.523	t_R=.563 t_D=.554
a=.1	SEX	t_R=.079 t_D=.071	t_R=.164 t_D=.138	t_R=.199 t_D=.169	t_R=.235 t_D=.210
	SER	t_R=.078 t_D=.074	t_R=.157 t_D=.139	t_R=.191 t_D=.166	t_R=.222 t_D=.210
	RULE	t_R=.082 t_D=.075	t_R=.164 t_D=.137	t_R=.200 t_D=.168	t_R=.228 t_D=.217

Table 3-5. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimator of σ_ε for Given a, b, N, and σ_ε. α = .05

N = 15 σ_ε = .03

		b=0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.039 t_D=.050	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000
	SER	t_R=.039 t_D=.039	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000
	RULE	t_R=.039 t_D=.050	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000
a=.3	SEX	t_R=.064 t_D=.034	t_R=1.000 t_D=.993	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000
	SER	t_R=.055 t_D=.029	t_R=.998 t_D=.985	t_R=1.000 t_D=.999	t_R=.999 t_D=.999
	RULE	t_R=.064 t_D=.034	t_R=1.000 t_D=.993	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000
a=.1	SEX	t_R=.060 t_D=.057	t_R=.625 t_D=.462	t_R=.694 t_D=.596	t_R=.637 t_D=.660
	SER	t_R=.056 t_D=.044	t_R=.598 t_D=.431	t_R=.654 t_D=.556	t_R=.587 t_D=.590
	RULE	t_R=.660 t_D=.057	t_R=.625 t_D=.462	t_R=.694 t_D=.596	t_R=.637 t_D=.660

Table 3-6. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimator of σ_ε for Given a, b, N, and σ_ε. α = .05

N = 15 σ_ε = .05

		b=0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.062 t_D=.041	t_R=1.000 t_D=.997	t_R=1.000 t_D=.999	t_R=1.000 t_D=1.000
	SER	t_R=.060 t_D=.033	t_R=.997 t_D=.991	t_R=.999 t_D=.998	t_R=.996 t_D=1.000
	RULE	t_R=.062 t_D=.041	t_R=1.000 t_D=.994	t_R=.999 t_D=.999	t_R=.999 t_D=1.000
a=.3	SEX	t_R=.062 t_D=.041	t_R=.948 t_D=.817	t_R=.976 t_D=.961	t_R=.966 t_D=.976
	SER	t_R=.060 t_D=.045	t_R=.930 t_D=.793	t_R=.958 t_D=.922	t_R=.895 t_D=.926
	RULE	t_R=.062 t_D=.044	t_R=.948 t_D=.815	t_R=.970 t_D=.955	t_R=.931 t_D=.967
a=.1	SEX	t_R=.053 t_D=.055	t_R=.327 t_D=.247	t_R=.361 t_D=.320	t_R=.356 t_D=.362
	SER	t_R=.054 t_D=.046	t_R=.304 t_D=.216	t_R=.340 t_D=.297	t_R=.331 t_D=.315
	RULE	t_R=.054 t_D=.054	t_R=.332 t_D=.253	t_R=.364 t_D=.326	t_R=.363 t_D=.364

Table 3-7. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimator of σ_ε for Given a, b, N, σ_ε. α = .05

N = 15 σ_ε = .07

		b=0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.055 t_D=.047	t_R=.990 t_D=.930	t_R=.998 t_D=.994	t_R=.994 t_D=.995
	SER	t_R=.047 t_D=.041	t_R=.984 t_D=.905	t_R=.991 t_D=.968	t_R=.953 t_D=.980
	RULE	t_R=.055 t_D=.047	t_R=.989 t_D=.920	t_R=.993 t_D=.973	t_R=.972 t_D=.982
a=.3	SEX	t_R=.064 t_D=.035	t_R=.755 t_D=.606	t_R=.848 t_D=.807	t_R=.828 t_D=.832
	SER	t_R=.059 t_D=.033	t_R=.728 t_D=.566	t_R=.780 t_D=.727	t_R=.741 t_D=.764
	RULE	t_R=.063 t_D=.036	t_R=.742 t_D=.589	t_R=.793 t_D=.744	t_R=.752 t_D=.781
a=.1	SEX	t_R=.052 t_D=.044	t_R=.217 t_D=.170	t_R=.242 t_D=.222	t_R=.233 t_D=.242
	SER	t_R=.047 t_D=.039	t_R=.213 t_D=.156	t_R=.237 t_D=.202	t_R=.216 t_D=.211
	RULE	t_R=.053 t_D=.043	t_R=.226 t_D=.173	t_R=.249 t_D=.214	t_R=.229 t_D=.235

Table 3-8. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimator of σ_ε for Given a, b, N, and σ_ε. α = .05

N=15 σ_ε = .09

		b=0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.058 t_D=.052	t_R=.910 t_D=.783	t_R=.976 t_D=.943	t_R=.930 t_D=.971
	SER	t_R=.045 t_D=.046	t_R=.879 t_D=.750	t_R=.933 t_D=.896	t_R=.850 t_D=.915
	RULE	t_R=.055 t_D=.052	t_R=.894 t_D=.779	t_R=.956 t_D=.913	t_R=.895 t_D=.934
a=.3	SEX	t_R=.045 t_D=.046	t_R=.597 t_D=.459	t_R=.710 t_D=.620	t_R=.634 t_D=.663
	SER	t_R=.041 t_D=.037	t_R=.590 t_D=.438	t_R=.657 t_D=.563	t_R=.574 t_D=.600
	RULE	t_R=.042 t_D=.040	t_R=.594 t_D=.447	t_R=.664 t_D=.567	t_R=.576 t_D=.604
a=.1	SEX	t_R=.051 t_D=.049	t_R=.165 t_D=.133	t_R=.171 t_D=.174	t_R=.174 t_D=.184
	SER	t_R=.052 t_D=.036	t_R=.160 t_D=.108	t_R=.161 t_D=.154	t_R=.164 t_D=.172
	RULE	t_R=.053 t_D=.040	t_R=.166 t_D=.115	t_R=.166 t_D=.158	t_R=.167 t_D=.175

Table 3-9. Percent of Significant Tests for Learning Using Three Different Methods for Choosing an Estimation of σ_ε for Given a, b, N, and σ_ε. $\alpha = .10$

$N = 6 \quad \sigma_\varepsilon = .03$

		b=.0	b=.4	b=.8	b=1.2
a=.5	SEX	t_R=.110 t_D=.094	t_R=1.000 t_D=.998	t_R=1.000 t_D=1.000	t_R=1.000 t_D=1.000
	SER	t_R=.105 t_D=.090	t_R=.999 t_D=.997	t_R=1.000 t_D=1.000	t_R=.999 t_D=1.000
	RULE	t_R=.110 t_D=.094	t_R=1.000 t_D=.998	t_R=1.000 t_D=1.000	t_R=.999 t_D=1.000
a=.3	SEX	t_R=.108 t_D=.090	t_R=.974 t_D=.952	t_R=.998 t_D=.995	t_R=.998 t_D=.998
	SER	t_R=.109 t_D=.089	t_R=.963 t_D=.929	t_R=.997 t_D=.990	t_R=.996 t_D=.993
	RULE	t_R=.109 t_D=.090	t_R=.972 t_D=.952	t_R=.998 t_D=.993	t_R=.996 t_D=.997
a=.1	SEX	t_R=.095 t_D=.083	t_R=.450 t_D=.404	t_R=.649 t_D=.594	t_R=.641 t_D=.635
	SER	t_R=.095 t_D=.075	t_R=.445 t_D=.399	t_R=.619 t_D=.573	t_R=.607 t_D=.596
	RULE	t_R=.095 t_D=.083	t_R=.450 t_D=.404	t_R=.649 t_D=.594	t_R=.641 t_D=.635

The primary concern is, however, to select the method for estimating σ_ε that will result in the most powerful test for detecting learning. Comparing the test procedure results with the estimator results it appears that using the minimum biased estimate of σ_ε does not necessarily result in the most powerful test (i.e. under those conditions when either equation (2-9) or the general rule presented on page 68 provided the largest percent of minimum biased estimates of σ_ε, the largest percent of significant test statistics is obtained when equation (2-8) is used to estimate σ_ε

Recall that when either parameter "a" or "b" is small, $(a \leq .1)$, $(b \leq .4)$, the observations are all very close to the asymptote. Under these conditions the rate of learning is very small and the corresponding average slope is small. If σ_ε is large, the probability of detecting learning is small. When σ_ε gets smaller, the probability of detecting learning increases. This same idea holds for estimates of σ_ε. Examining the distribution of the estimates for both estimators in Figures 3-15 through 3-22 it appears that equation (2-8) will always produce the larger percent of smaller estimates of σ_ε. Therefore using estimates of σ_ε generated using equation (2-8) when applying the LLSR test procedure for any combination of sample size, σ_ε and parameter values results in a more powerful test for learning.

Table 3-10. Results for Each of the Three Methods for Choosing an Estimator of σ_ε in Terms of Percent of Minimum Biased Estimates of σ_ε

$N = 6 \quad \sigma_\varepsilon = .03$

	b=0	b=.4	b=.8	b=1.2
	SEX=.445	SEX=.776	SEX=.988	SEX=.999
a=.5	SER=.555	SER=.224	SER=.012	SER=.001
	RULE=.447	RULE=.778	RULE=.897	RULE=.583
	SEX=.442	SEX=.548	SEX=.823	SEX=.948
a=.3	SER=.558	SER=.452	SER=.177	SER=.052
	RULE=.552	RULE=.564	RULE=.775	RULE=.735
	SEX=.455	SEX=.440	SEX=.471	SEX=.514
a=.1	SER=.545	SER=.560	SER=.529	SER=.406
	RULE=.461	RULE=.459	RULE=.476	RULE=.526

Table 3-11. Results for Each of the Three Methods for Choosing an Estimator of σ_ε in Terms of Percent of Minimum Biased Estimates of σ_ε

$N = 6 \quad \sigma_\varepsilon = .05$

	b=0	b=.4	b=.8	b=1.2
	SEX=.576	SEX=.560	SEX=.794	SEX=.952
a=.5	SER=.424	SER=.440	SER=.206	SER=.048
	RULE=.486	RULE=.584	RULE=.667	RULE=.593
	SEX=.471	SEX=.477	SEX=.569	SEX=.714
a=.3	SER=.528	SER=.523	SER=.431	SER=.286
	RULE=.549	RULE=.520	RULE=.470	RULE=.397
	SEX=.431	SEX=.435	SEX=.492	SEX=.479
a=.1	SER=.569	SER=.565	SER=.518	SER=.521
	RULE=.536	RULE=.548	RULE=.519	RULE=.533

Table 3-12. Results for Each of the Three Methods for Choosing an Estimator of σ_ε in Terms of Percent of Minimum Biased Estimates of σ_ε

N = 6 σ_ε = .07

	b=0	b=.4	b=.8	b=1.2
	SEX=.543	SEX=.487	SEX=.741	SEX=.798
a=.5	SER=.457	SER=.513	SER=.256	SER=.202
	RULE=.473	RULE=.517	RULE=.539	RULE=.547
	SEX=.439	SEX=.467	SEX=.530	SEX=.556
a=.3	SER=.561	SER=.533	SER=.470	SER=.444
	RULE=.458	RULE=.443	RULE=.380	RULE=.317
	SEX=.461	SEX=.438	SEX=.456	SEX=.452
a=.1	SER=.539	SER=.562	SER=.544	SER=.548
	RULE=.448	RULE=.459	RULE=.465	RULE=.445

Table 3-13. Results for Each of the Three Methods for Choosing an Estimator of σ_ε in Terms of Percent of Minimum Biased Estimates of σ_ε

N = 6 σ_ε = .09

	b=0	b=.4	b=.8	b=1.2
	SEX=.449	SEX=.498	SEX=.558	SEX=.668
a=.5	SER=.551	SER=.502	SER=.442	SER=.332
	RULE=.443	RULE=.473	RULE=.554	RULE=.601
	SEX=.453	SEX=.467	SEX=.467	SEX=.504
a=.3	SER=.547	SER=.533	SER=.533	SER=.496
	RULE=.492	RULE=.482	RULE=.438	RULE=.447
	SEX=.440	SEX=.454	SEX=.444	SEX=.477
a=.1	SER=.560	SER=.540	SER=.556	SER=.532
	RULE=.498	RULE=.472	RULE=.494	RULE=.472

Table 3-14. Results for Each of the Three Methods for Choosing an Estimator of σ_ε in Terms of Percent of Minimum Biased Estimates of σ_ε

$N = 15 \qquad \sigma_\varepsilon = .03$

	b=0	b=.4	b=.8	b=1.2
	SEX=.426	SEX=.991	SEX=.999	SEX=1.000
a=.5	SER=.574	SER=.009	SER=.001	SER=0.00
	RULE=.426	RULE=.989	RULE=.995	RULE=.893
	SEX=.411	SEX=.788	SEX=.991	SEX=.999
a=.3	SER=.589	SER=.212	SER=.009	SER=.001
	RULE=.411	RULE=.788	RULE=.992	RULE=.991
	SEX=.397	SEX=.418	SEX=.506	SEX=.555
a=.1	SER=.603	SER=.582	SER=.494	SER=.445
	RULE=.397	RULE=.445	RULE=.475	RULE=.572

Table 3-15. Results for Each of the Three Methods for Choosing an Estimator of σ_ε in Terms of Percent of Minimum Biased Estimates of σ_ε

$N = 15 \qquad \sigma_\varepsilon = .05$

	b=0	b=.4	b=.8	b=1.2
	SEX=.426	SEX=.784	SEX=.991	SEX=1.000
a=.5	SER=.574	SER=.216	SER=.009	SER=0.000
	RULE=.487	RULE=.734	RULE=.794	RULE=.644
	SEX=.387	SEX=.509	SEX=.750	SEX=.896
a=.3	SER=.613	SER=.491	SER=.250	SER=.104
	RULE=.508	RULE=.555	RULE=.586	RULE=.542
	SEX=.501	SEX=.422	SEX=.436	SEX=.474
a=.1	SER=.499	SER=.578	SER=.564	SER=.526
	RULE=.501	RULE=.490	RULE=.506	RULE=.506

Table 3-16. Results for Each of the The Three Methods for Choosing an Estimator of σ_ε in Terms of Percent of Minimum Biased Estimates of σ_ε

N = 15 σ_ε = .07

	b=0	b=.4	b=.8	b=1.2
	SEX=.405	SEX=.612	SEX=.893	SEX=.969
a=.5	SER=.595	SER=.388	SER=.107	SER=.031
	RULE=.430	RULE=.526	RULE=.568	RULE=.517
	SEX=.434	SEX=.467	SEX=.597	SEX=.662
a=.3	SER=.566	SER=.533	SER=.403	SER=.338
	RULE=.469	RULE=.370	RULE=.263	RULE=.216
	SEX=.443	SEX=.411	SEX=.440	SEX=.419
a=.1	SER=.557	SER=.589	SER=.560	SER=.581
	RULE=.412	RULE=.449	RULE=.411	RULE=.394

Table 3-17. Results for Each of the Three Methods for Choosing an Estimator of σ_ε in Terms of Percent of Minimum Biased Estimates of σ_ε

N = 15 σ_ε = .09

	b=0	b=.4	b=.8	b=1.2
	SEX=.427	SEX=.546	SEX=.736	SEX=.853
a=.5	SER=.573	SER=.454	SER=.264	SER=.147
	RULE=.490	RULE=.569	RULE=.644	RULE=.620
	SEX=.422	SEX=.456	SEX=.474	SEX=.567
a=.3	SER=.578	SER=.544	SER=.526	SER=.433
	RULE=.553	RULE=.516	RULE=.478	RULE=.438
	SEX=.429	SEX=.440	SEX=.440	SEX=.421
a=.1	SER=.571	SER=.560	SER=.560	SER=.579
	RULE=.526	RULE=.546	RULE=.526	RULE=.563

Evaluation of the Nonlinear Test Procedure

A final simulation study was conducted to determine if the nonlinear test procedure is a more powerful test than the linear test procedures. The study was conducted under exactly the same conditions as were used for the linear test evaluation (i.e. same normal randomly generated observations were used). In order for the nonlinear test procedure to yield a significant test statistic, two conditions must be satisfied:

(i) The degree of nonlinearity of the curve must be small enough to apply linear theory approximation for determining confidence limits on parameters "a" and "b".

(ii) The lower limit for both parameter "a" and parameter "b" must be greater than zero for the lower confidence limit of the slope abt^{-b-1} to be greater than zero.

An advantage of the nonlinear test over the linear test procedure is that it is possible to estimate the rate of learning at any particular trial by examining the confidence limits for the true slope at that trial. Therefore if the degree of nonlinearity is small enough, less than $.01/F_{\alpha,p,N-p}$, it will be possible to construct the upper $(1-\alpha)$ confidence limit for the true rate of learning at any particular trial. If the value of the upper confidence limit, in terms of rate of learning is determined to be insignificant at a particular trial then we can assume that learning will not be a factor in any future trial results.

As can be seen in Table 3-18 when either parameter "a" decreases or parameter "b" increases, the maximum allowable σ_ε value for which linear theory approximations on the test of the slope will be valid, decreases. For the nonlinear test procedure, the results in Tables 3-20 and 3-21 indicate that the degree of nonlinearity is the limiting factor except when parameter "a" or "b" is small. An explanation for this follows. When parameter "b" is small the rate of learning is small and the curve is approximately a straight line. A tangent line approximation for estimating the values of the parameters would do very well under these conditions. When σ_ε is large in comparison to the average rate of learning, it will reduce the power of the test before it becomes a significant factor in affecting the degree of nonlinearity. If on the other hand the value of parameter "a" is small, this implies that the observations are all very near the asymptotic value of 1. Since we can never do any better than 100%, the amount of deviation above the expected value is limited. If σ_ε is relatively large the amount of deviation that occurs below the expected value may well exceed the limited deviation above this expected value. This would result in a fitted curve with a negative value for b, but more important it will also result in a lower confidence limit for the slope which includes zero. An examination of the observations for several simulation runs when parameter a = .1 and σ_ε = .03 did reveal this to be the case.

Table 3-18. Maximum Allowable Value of σ_ε for Given Values of "a" and "b" for Linear Theory Approximations to be Valid at α = .05 Level

N = 6

	b=0.0	b=.4	b=.8	b=1.2
a=.5	∞	.078997	.042525	.031026
a=.3	∞	.047398	.025515	.018615
a=.1	∞	.015799	.008505	.006205

Table 3-19. Maximum Allowable Value of σ_ε for Given Values of "a" and "b" for Linear Theory Approximations to be Valid at α = .05 Level

N = 15

	b=0	b=.4	b=.8	b=1.2
a=.1	∞	.073523	.035870	.025137
a=.3	∞	.044114	.021523	.015022
a=.5	∞	.014705	.007174	.005027

Table 3-20. Percent of Times Linear Approximations can be Used, P_{LA} and the Percent of Times the Slope Test was Significant, P_{SS}

N = 6 σ_ε = .03

	b=0	b=.4	b=.8	b=1.2
a = .5	P_{LA}=1.000	P_{LA}=1.000	P_{LA}=.895	P_{LA}=.623
	P_{SS}= .080	P_{SS}=1.000	P_{SS}=.895	P_{SS}=.623
a = .3	P_{LA}=1.000	P_{LA}= .956	P_{LA}=.399	P_{LA}=.200
	P_{SS}= .071	P_{SS}= .938	P_{SS}=.399	P_{SS}=.200
a = .1	P_{LA}= .794	P_{LA}= .267	P_{LA}=.056	P_{LA}=.041
	P_{SS}= .053	P_{SS}= .079	P_{SS}=.025	P_{SS}=.025

Table 3-21. Percent of Times Linear Approximations can be Used P_{LA} and the Percent of Times the Slope Test was Significant P_{SS}

N = 6 σ_ε = .05

	b=0	b=.4	b=.8	b=1.2
a = .5	P_{LA}=1.000	P_{LA}= .951	P_{LA}=.438	P_{LA}=.187
	P_{SS}= .080	P_{SS}= .940	P_{SS}=.438	P_{SS}=.187
a = .3	P_{LA}= .994	P_{LA}= .579	P_{LA}=.104	P_{LA}=.056
	P_{SS}= .072	P_{SS}= .439	P_{SS}=.103	P_{SS}=.056
a = .1	P_{LA}= .145	P_{LA}= .154	P_{LA}=.065	P_{LA}=.059
	P_{SS}= .015	P_{SS}= .008	P_{SS}=.005	P_{SS}=.016

The results in terms of the percent of significant tests per 1000 runs using the nonlinear test procedure are given in Tables 3-22 and 3-23. As the value of σ_ε decreases, the power of the nonlinear test increases when the values of parameters a and b exceed some critical number. It appears that this critical value may be near .2. As the sample size increases the power of the nonlinear test decreases. This is evident by comparing values in Tables 3-18 and 3-19 since the maximum value of σ_ε for given parameter values "a" and "b" decreases as N increases.

A comparison of the LLSR test procedure with the nonlinear test procedure is also provided in Tables (3-22) and (3-23).

Table 3-22. Comparison of the Percent of Significant Tests for Learning Using the Nonlinear Procedure, t_{NL}, and the LLSR Procedure, t_R. The results are based on 1000 simulation runs for each combination of a, b, N, and σ_ε. Tests were conducted at α = .05 level.

N = 6 σ_ε = .03

	b=0	b=.4	b=.8	b=1.2
a=.5	t_{NL}=.080 t_R =.067	t_{NL}=1.000 t_R = .998	t_{NL}= .895 t_R =1.000	t_{NL}= .623 t_R =1.000
a=.3	t_{NL}=.071 t_R =.059	t_{NL}= .938 t_R = .913	t_{NL}= .399 t_R = .983	t_{NL}= .200 t_R = .982
a=.1	t_{NL}=.053 t_R =.056	t_{NL}= .079 t_R = .318	t_{NL}= .025 t_R = .415	t_{NL}= .025 t_R = .497

Table 3-23. Comparison of the Percent of Significant Tests for Learning Using the Nonlinear Procedure, t_{NL}, and the LLSR Procedure, t_R. The results are based on 1000 simulation runs for each combination of a, b, N and σ_ε. Tests were conducted at the α = .05 level.

N = 6 σ_ε = .05

	b=0	b=.4	b=.8	b=1.2
a=.5	t_{NL}=.080 t_R =.086	t_{NL}= .940 t_R = .958	t_{NL}= .438 t_R = .997	t_{NL}= .187 t_R = .995
a=.3	t_{NL}=.072 t_R =.080	t_{NL}= .439 t_R = .715	t_{NL}= .103 t_R = .912	t_{NL}= .056 t_R = .894
a=.1	t_{NL}=.015 t_R =.084	t_{NL}= .008 t_R = .257	t_{NL}= .005 t_R = .328	t_{NL}= .016 t_R = .381

CHAPTER IV

ILLUSTRATION OF THE PROPOSED PROCEDURE

Example

The example problem chosen was an actual experiment conducted at Georgia Tech to determine the performance of a viscous damped tripod. The experimental results were in terms of standard deviation of the error from a marked point while tracking a moving target at a constant velocity. The curve describing these results follows the form of equation (1-1). Although the linear test for learning can be applied directly to the learning curve data, a conversion to the performance curve will be made in order to apply the computer program as writtin for the nonlinear test procedure.

To conduct a test for learning using the performance curve described by equation (1-2), a suitable scaling of the values must be accomplished. Recall when performance is in terms of the percent of total possible that the lower limit at trial 1 was 0 and the upper limit of the curve at some future trial number was 1. A value of zero is what we might expect from an operator who is totally unfamiliar with the system and a value of 100% is the expected value obtained if the system meets the required specifications when operated by a fully trained individual or crew.

In the viscous damped tripod experimental data, values that correspond to the lower and upper limits for the performance as discussed above must be selected. The upper limit will be chosen as the minimum standard deviation of error predicted by the manufacturer when the tripod is operated by a fully learned individual. The value selected to correspond to the lower limit on the performance curve was the largest standard deviation of error value recorded during the first two trials of the experiment. The tripod was operated by 5 different individuals whose experience in tracking moving targets varied. The largest standard deviation of error over the first two trials was recorded by a subject who was totally naive about the operation of the tripod prior to this experiment.

The equation for obtaining the proper scaling factor to use in transforming the learning curve data to a performance curve is:

$$T_F = \frac{1.0}{E_S - E_M} \tag{4-1}$$

where

E_M represents the manufacturer's specifications

E_S represents the largest value for standard deviation of error recorded during the first two trials

T_F = transformation factor for 1 unit of change in the learning curve data.

The interpretation of the above computations is that 1 unit of decrease from the largest value recorded in the learning curve data will correspond to $1.0/[E_S - E_M]$ increase from the minimum value, zero, on the performance curve. The equation for transforming a data point at a particular trial on the learning curve to the performance curve is

$$V_p = (E_S - \text{trial result})\ (T_F) \qquad (4\text{-}2)$$

$$V_p = (6.522 - \text{trial result})(.18443)$$

where V_p represents the corresponding value on the performance curve.

The results in column 2 below were obtained by an individual who was familiar with tracking moving objects but who had never operated this particular type of tripod before this experiment. A test to detect if learning was occurring during the first 6 trials of the experiment will be conducted using both the LLSR test procedure and the nonlinear test procedure. If the nonlinear test is significant, the 100 $(1-\alpha)$% confidence limits will be evaluated for the slope of the curve at each trial to determine if the rate of learning becomes insignificant by trial 6.

Trial Number	Result	$V_p = y_i$	X_i	$(X_i - \bar{X})^2$	$(t_i - \bar{t})y_i$
1	4.2124	.4260	.1328	.003564	-1.0650
2	3.4920	.5588	.0778	.000022	- .8382
3	3.0702	.6366	.1582	.007242	- .3183
4	2.2126	.7948	.0740	.000001	.3974
5	1.8113	.8688	-.0773	.022620	1.3032
6	2.2306	.7915			1.9789

Applying the LLSR test procedure at $\alpha = .05$ using an estimate of σ_ε from equation (2-8) yields the following:

$$H_0: \hat{d}_{LLSR} \leq 0$$

$$H_0: \hat{d}_{LLSR} > 0$$

Compute:

$$t_0 = \frac{\hat{d}_{LLSR} - 0}{\sqrt{\dfrac{12\hat{\sigma}_\varepsilon^2}{N(N^2+1)}}}$$

If $t_0 \leq t_{.05,4}$ do not reject H_0

If $t_0 > t_{.05,4}$ reject H_0

Compute an estimate of σ_ε^2 using equation (2-8)

$$\hat{\sigma}_\varepsilon^2 = \frac{(N-1) \sum_{i=1}^{5} (x_i - \bar{x})^2}{2N(N-2)}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{5(.033449)}{2(6)(4)} = .0034843$$

Compute an estimate of the slope using the LLSR method:

$$\hat{d}_{LLSR} = \frac{\sum_{i=1}^{6} (t_i - \bar{t}) y_i}{\sum_{i=1}^{N} (t_i - \bar{t})^2}$$

$$\hat{d}_{LLSR} = \frac{1.458}{17.5}$$

$$\hat{d}_{LLSR} = .0833$$

Compute the test statistic, t_o:

$$t_o = \frac{\hat{d}_{LLSR} - 0}{\sqrt{\frac{12\sigma_\varepsilon^2}{N(N^2+1)}}}$$

$$t_o = \frac{.0833}{\sqrt{\frac{12(.003484}{6(36+1)}}}$$

$$t_o = 6.07$$

Since $t_o > t_{05,4}$ we reject H_0 and conclude learning is occurring during these 6 trials.

Applying the nonlinear test procedure at $\alpha = .05$ using the computer program in Appendix D, yields the following results.

Estimates of the parameters and σ_ε are:

$\hat{a} = .6017$

$\hat{b} = .6425$

$\hat{\sigma}_\varepsilon = .0656$

The measure of non linearity $\hat{N}_\theta$ was computed as:

$\hat{N}_\theta = .00161373$

Critical Value = .001441

The degree of nonlinearity for this estimated curve is too large to use linear theory approximations to compute 95% confidence limits on the slope. Applying the test procedure at $\alpha = .10$, the test for learning is significant. The 90% confidence interval on the slope at each trial is:

Trial Number	
1	lower confidence value = .1769 upper confidence value = .7366
2	lower confidence value = .0661 upper confidence value = .1902
3	lower confidence value = .0372 upper confidence value = .0878
4	lower confidence value = .0247 upper confidence value = .0518

5 lower confidence value = .0180
upper confidence value = .0348

6 lower confidence value = .0139
upper confidence value = .0254

This particular operator, even though familiar with tracking targets, still appears to be learning after 6 trials. Therefore we would conclude that none of the 6 trial observation values are representative of the performance of this tripod when operated by a fully learned individual.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

This research involved the development of a simple methodology to test for learning in experimental results using small sample sizes. In addition, a procedure for examining the instantaneous rate of learning at any particular trial of an experiment was also investigated.

Assuming learning can be described by a monotonically increasing performance curve of the form $z = 1\text{-}at^{-b}$, tests for learning were developed based on examining the rate of learning over several trails. Since the curve is monotonically increasing, a positive slope will be interpreted as learning and a zero slope will correspond to no learning occurring. For this research the time between trials was considered insignificant in affecting previously gained knowledge and the error between any observation and its expected value, z_i, is assumed to be NID $(0, \sigma_\varepsilon^2)$.

To develop a simple methodology to test for learning and to provide a test that would measure the rate of learning at any particular trial, required two different approaches to the problem. The approach to developing a simplified test procedure involved examining four linear methods used to

estimate the average rate of learning over several trials. A comparison of the variances of the estimates of the average rate of learning for each method was made and those estimates with minimum variance were sleected for further analysis. The best linear method was the linear least squares regression, LLSR, method and the next best linear method was the average of the consecutive differences, ACD, method. Although the variance of LLSR estimate of the average rate of learning is smaller than the variance of the ACD method, the expected value of the LLSR estimate contains a negative bias factor while the expected value of the estimate of the average of learning using the ACD method is unbiased. The amount of bias in the LLSR estimate of the average rate of learning increases as the sample size or the initial rate of learning increases. When the ratio of the LLSR estimate of the average rate of learning to the ACD estimate of average rate of learning is smaller than the ratio of their corresponding standard deviations, then the ACD method becomes the more powerful test procedure. For sample sizes less that 24, the LLSR method is the best test procedure when a parameter b is less than or equal to one.

In order to conduct either of the linear test procedures an estimate of the process variance, σ_ε^2, must be obtained. Three estimators of σ_ε^2 were examined to determine which

estimator provided estimators of σ_ε^2 which resulted in the most powerful test for learning. These estimators were derived using the LLSR estimate of σ_ε^2, the sum of the squared differences between the observations and their average, and the sum of the square differneces between the first difference of the observations and their average. The estimator of σ_ε^2 using the first difference of the observations provided estimates which resulted in the most powerful tests for learning.

Since the rate of learning in performance evaluations is not expected to exceed a parameter value of $b > 1.0$, the most powerful test procedure would be the LLSR. This method is most powerful when the variance of the slope estimate is computed with an estimate of σ_ε^2 obtained using the first difference estimator. To measure the rate of learning at any particular trial requires an analysis of the instantaneous slope of the curve at that point. The approach taken was to use linear theory approximations to estimate the slope. If the degree of nonlinearity of the function is small enough, it is possible to use a linear theory approximation to construct a confidence interval for the true slope at any particular trial. In the nonlinear method, estimates for σ_ε and the parameters "a" and "b" are obtained and a test on the degree of nonlinearity of the function is conducted using Beale's measure of nonlinearity. If the degree of nonlinearity is small enough then it is possible to construct a confidence region for the parameters "a" and "b". By

computing the estimate of the slope over the points on the periphery of the parameter confidence region, the smallest and the largest value of the slope estimate can be obtained for the curve at a given trial. The smallest and largest value of the slope estimate correspond to the lower and upper confidence limits respectively on the true slope. The width of the confidence interval for the true slope will decrease over the latter trials until it will eventually include the value zero which indicates a fully learned status. Therefore if the upper confidence limit were less than some maximum acceptable rate of learning at a certain trial, then it would be concluded that learning would not be a factor in any future trial results.

In a comparison of the two procedures, the linear methods were more powerful tests; however, the nonlinear method was able to provide information on the rate of learning at each trial when the nonlinearity conditions were satisfied and significant learning was detected. When the degree of nonlinearity was small enough to conduct the test, significant learning was detected 95% of the time except when parameter "a" and "b" were small. The more powerful linear test procedure was the LLSR method, which can detect an average rate of learning over 15 trials of .01 at an $\alpha = .05$ level 95% of the time when the standard deviation is $\sigma_\varepsilon \leq .05$.

The advantages of the linear test procedures over the nonlinear procedure are greater probability of detecting actual learning and the test can be applied using simple arithmetic and comparing computed and tabulated values as demonstrated with an example in Chapter IV.

Recommendations for Future Study

In order to apply the nonlinear test procedure on results which follow the learning curve, equation (1-1), the data must be transformed into data that follows the form of a performance curve, equation (1-2). The critical factor in the transformation of data is selecting a value from the learning curve data that would correspond to zero performance on the performance curve. If the value selected is too small, the scaling factor, T_F, will be inflated and when the value selected is too large T_F will be deflated. The effect on the power of the test procedures is not clear since a deflated T_F results in a larger estimate for σ_ε^2 as well as a larger estimate for the average rate of learning. It is recommended that future research be conducted to investigate the effect of errors in this transformation.

If the error term in the learning curve model was multiplicative rather than additive as assumed in this study, the test procedures developed may not be adequate for detecting learning. It may be necessary to develop a new test procedure based on the logarithmic transform of the model,

$$\ln(y) = \ln(a) + b\ln(t_i) + \ln(\varepsilon_i)$$

where $\ln(y)$ = logarithm (y)

A recommendation for future research would be to develop a methodology for detecting learning when the errors are log normally distributed. Assuming that $\ln(y_i)$ is normally distributed then $\ln(a)$ and b will be normally distributed; however, the distribution of $e^{\ln(a)} = a$ is not normally distributed. The difficulty then in developing a test procedure for this approach will be in determining the distribution of the estimate of the slope, $-abt^{-b-1}$.

If the value of parameter b is known, it would be possible to select the test procedure which would be more powerful for detecting learning. It is known that if the value of parameter b increases for a given sample size and and a given value for "a", that the average rate of learning also increases. If an estimate of the value of "a" is made using the first observation and the value of N is known, then it may be possible to determine the value of b by examining the estimate of the average rate of learning. Another recommendation for future research is to study the relationship between the average rate of learning and parameter b in order to increase the probability of choosing the more powerful test when the true performance curve is unknown.

APPENDIX A

EXPLANATION OF NOTATION

Chapter I

z_i	value of the true performance curve at trial number i
σ_ε^2	variance of an observation about its expected value
ε_i	error between the observation and its expected value at trial i

Chapter II

y_i	the observation at trial number i
z_c	value of the performance curve at the asymtote
$\bar{y}$	average of the observations
x_i	the difference between y_{i+1} and y_i
$\bar{x}$	average of the consecutive differences between the observations
ℓ	average rate of learning
$\hat{d}_{ACD}$	estimate of the average rate of learning using the average of consecutive differences (ACD) method
$\hat{d}_{LLSR}$	estimate of the average rate of learning using the the Linear Lease Squares Regression (LLSR) method
S_V^2	minimum variance unbiased estimate of σ_ε^2
$\hat{N}_\theta$	measure of degree of nonlinearity
$\eta_i(\bar{\theta})$	true value of the function at trial i when the parameter values are $\bar{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$
$\tau_i(\bar{\theta})$	tangential approximation of the true function at trial i when the parameter values are $\bar{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$
$(OBS)^2$	estimate of σ_ε^2 using equation (2-7)

$(SEX)^2$ estimate of σ_ε^2 using equation (2-8)

$(SER)^2$ estimate of σ_ε^2 using equation (2-9)

Chapter III

$E[\hat{\sigma}_{\varepsilon_X}]$ expected value of the estimate of σ_ε using the average of consecutive differences estimator, equation (2-8)

$E[\hat{\sigma}_{\varepsilon_R}]$ expected value of the estimate of σ_ε using the linear least squares regression estimator (equation 2-9)

OBS estimate of σ_ε using equation (2-7)

SEX estimate of σ_ε using equation (2-8)

SER estimate of σ_ε using equation (2-9)

RULE estimate of σ_ε applying the general rule described on page 68 for choosing an estimator

t_R percent of times the Linear Least Squares Regression test procedure detected significant learning

t_D percent of times the Average of Consecutive Differences test procedure detected significant learning

P_{LA} percent of times that the degree of nonlinearity of the performance curve was small enough to use linear theory approximations

P_{SS} percent of times the nonlinear test procedure detected significant learning when linear theory approximations could be applied.

t_{NL} percent of time the nonlinear test procedure detected significant learning

T_F scaling factor for transforming learning curve data to a performance curve

E_S largest value recorded in the learning curve data

APPENDIX B

DERIVATION OF THE ESTIMATORS OF σ_ε^2

Several alternative techniques for estimating the variance, σ_ε^2, are examined in terms of expected bias to dtermine a minimum bias estimator. The first technique considered will be that presented in equation (2-6).

$$S_\varepsilon^2 = \sum_{i=1}^{N} (y_i - \bar{y})^2/(N-1) \qquad \text{(B-1)}$$

writing S_ε^2 as a function of the z_i's and taking expectations we have

$$E(S_\varepsilon^2) = E[\sum_{i=1}^{N} (y_i - \bar{y})/(N-1)]$$

$$= \frac{1}{N-1} E\{\sum_{i=1}^{N} (y_i^2 - 2y_i\bar{y} + \bar{y}^2)$$

$$= \frac{1}{N-1}[E \sum_{i=1}^{N} \{(z_i+\varepsilon_i)^2 - 2(z_i+\varepsilon_i)(\frac{\sum_{i=1}^{N}(z_i+\varepsilon_i)}{N}) + (\sum_{i=1}^{N} \frac{(z_i+\varepsilon_i)}{N})^2\}$$

Since z_i is a constant and $\varepsilon_i \sim N10\ (0,\sigma_\varepsilon^{\ 2})$ the equation in reduced form is

$$E(S_\varepsilon^{\ 2}) = \frac{1}{N-1}\Big[\sum_{i=1}^{N}(z_i^{\ 2}+E(\varepsilon_i)^2)-2\sum_{i=1}^{N}\Big(\frac{z_i\sum_{i=1}^{N} z_i}{N}\Big)-2\sum_{i=1}^{N}\Big[\varepsilon_i\Big(\frac{\sum_{i=1}^{N} i}{N}\Big)\Big]+\sum_{i=1}^{N}E\Big(\frac{\sum_{i=1}^{N} z_i}{N}\Big)^2+\sum_{i=1}^{N}E\Big(\frac{\sum_{i=1}^{N}\varepsilon_i}{N}\Big)^2\Big]$$

Simplifying where $E(\varepsilon_i^{\ 2}) = \sigma_\varepsilon^{\ 2}$ and $E\Big(\frac{\sum_{i=1}^{N}\varepsilon_i}{N}\Big)^2 = \sigma_\varepsilon^{\ 2}/N$

$$E(S_\varepsilon^{\ 2}) = \sum_{i=1}^{N} z_i^{\ 2}/N-1 + \frac{N\sigma_\varepsilon^{\ 2}}{N-1} - \frac{2\sum_{i=1}^{N} z_i\bar{z}}{N-1}$$

$$- \frac{2N\ \sigma_\varepsilon^{\ 2}/N}{N-1} + \frac{\sum_{i=1}^{N}\bar{z}^2}{N-1} + \frac{N\sigma_\varepsilon^{\ 2}}{N(N-1)}$$

where $\bar{z}$ is defined as the average of the expected values of the observations.

$$E(S_\varepsilon^{\ 2}) = \sigma_\varepsilon^{\ 2} + \frac{1}{N-1}\Big[\sum_{i=1}^{N} z_i^{\ 2}-2\sum_{i=1}^{N} z_i\bar{z} + N\bar{z}^2\Big]$$

$$E(S_\varepsilon^{\ 2}) = \sigma_\varepsilon^{\ 2} + \frac{1}{N-1}\sum_{i=1}^{N}[z_i-\bar{z}]^2 \qquad \text{(B-2)}$$

Thus the farther the observations or trial results are from the asymptote the more inflated will be our estimate of the true error variance. A candidate for consideration as an estimator for the variance then is equation (B-1).

Another simple to apply method for estimating the variance of the process would be to use the first difference of the series of observations

$$x_i = y_{i+1} - y_i$$

where y_i is the observation at trial i.
Although the differences are not independent, the variance of the x_i's will contain the term σ_ε^2. Therefore it may be possible to express the variance of the x_i's as:

$$A \text{ Var } (x_i) = \sigma_\varepsilon^2 + B \qquad \text{(B-3)}$$

where A defines some multiplier and B represents the bias factor due to lack of fit of the model and dependence between the x_i values. Then a comparison of estimators can be made in terms of expected bias. The expected bias in the estimate of σ_ε^2 using the sum of squared errors of the consecutive differences between observations about the average

$$S_x^2 = \sum_{i=1}^{N-1} (x_i - \bar{x})^2 / N-2$$

is derived below

$$E(S_x^2) = E[\frac{\Sigma(x_i - \bar{x})^2}{N-2}]$$

$$E[\frac{\sum_{i=1}^{N-1}(x_i - \bar{x})^2}{N-2}] = \frac{1}{N-2} E\{\sum_{i=1}^{N-1}[(y_{i+1} - y_i) - \frac{\sum_{i=1}^{N-1}(y_{i+1} - y_i)}{N-1}]^2$$

$$= \frac{1}{N-2} E\{\sum_{i=1}^{N-1}[\{(z_{i+1} + \varepsilon_{i+1}) - (z_i + \varepsilon_i)\} - \frac{\sum_{i=1}^{N-1}\{(z_{i+1} + \varepsilon_{i+1}) - (z_i + \varepsilon_i)\}}{N-1}]^2$$

$$= \frac{1}{N-2} E\{\sum_{i=1}^{N-1}[(z_{i+1} + \varepsilon_{i+1} - z_i - \varepsilon_i)^2 - 2\{(z_{i+1} + \varepsilon_{i+1} - z_i - \varepsilon_i)$$

$$(\frac{\sum_{i=1}^{N-1}[z_{i+1} + \varepsilon_{i+1} - z_i - \varepsilon_i]}{N-1}) + (\frac{\sum_{i=1}^{N-1}[z_{i+1} + \varepsilon_{i+1} - z_i - \varepsilon_i]}{N-1})^2]\}$$

$$E[\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-2}] = \frac{1}{N-2}\{E\sum_{i=1}^{N-1}(z_{I=1} + \varepsilon_{i+1} - z_i - \varepsilon_i)^2 - 2E\sum_{i=1}^{N-1} \qquad \text{(B-4)}$$

$$\{(z_{i+1} + \varepsilon_{i+1} - z_i - \varepsilon_i)(\frac{\sum_{i=1}^{N}[z_{i+1} + \varepsilon_{i+1} - z_i - \varepsilon_i]}{N-1})\} + E\sum_{i=1}^{N-1}$$

$$(\frac{\sum_{i=1}^{N-1}[z_{i+1} + \varepsilon_{i+1} - z_i - \varepsilon_i]}{N-1})^2$$

The right side of equation (B-4) may be considered as separate terms and the expected value of each computed.

First term:

$$E\{\sum_{i=1}^{N-1}(z_{i+1}+\varepsilon_{i+1}-z_i-\varepsilon_i)^2\} = z_1^2+2\sum_{i=1}^{N-1}z_i^2+z_N^2-2\sum_{i=1}^{N-1}z_iz_{i+1}$$

$$+2(N-1)\sigma_\varepsilon^2$$

$$= \sum_{i=1}^{N-1}(z_{i+1}-z_i)^2 + 2(N-1)\sigma_\varepsilon^2$$

Second term:

$$-2E\{\sum_{i=1}^{N-1}[(z_{i+1}+\varepsilon_{i+1}-z_i-\varepsilon_i)(\frac{\sum_{i=1}^{N-1}[z_{i+1}+\varepsilon_{i+1}-z_i-\varepsilon_i]}{N-1})]\}$$

$$= -2E\{\sum_{i=1}^{N-1}[(z_{i+1}+\varepsilon_{i+1}-z_i-\varepsilon_i)(\frac{z_N+\varepsilon_N-z_1-\varepsilon_1}{N-1})]\}$$

$$= -2E\{(\frac{z_N+\varepsilon_N-z_1-\varepsilon_1}{N-1})\sum_{i=1}^{N-1}(z_{i+1}+\varepsilon_{i+1}-z_1-\varepsilon_1)\}$$

$$= -2E\{(\frac{z_N+\varepsilon_N-z_1-\varepsilon_1}{N-1})(z_N+\varepsilon_N-z_1-\varepsilon_1)\}$$

$$= -\frac{2}{N-1}E\{z_N^2+z_1^2-2z_Nz_1+2z_N\varepsilon_N-2z_N\varepsilon_1-2z_1\varepsilon_N-2\varepsilon_N\varepsilon_1+2z_1\varepsilon_1+\varepsilon_N^2+\varepsilon_1^2$$

Taking the expected value of each quantity inside the parentheses the following result is obtained

$$= -\frac{2}{N-1}(z_N^2+z_1^2-2z_Nz_1+2\sigma_\varepsilon^2)$$

$$= -\frac{2}{N-1}(z_N-z_1)^2 - \frac{4\sigma_\varepsilon^2}{N-1}$$

Third term:

$$E\{\sum_{i=1}^{N-1}[\frac{\sum_{i=1}^{N-1}(z_{i+1}+\varepsilon_i-z_i-\varepsilon_i)}{N-1}]^2\} = \sum_{i=1}^{N-1}[\frac{(z_N+\varepsilon_N-z_1+\varepsilon_1)}{N-1}]^2$$

$$= \frac{1}{(N-1)^2}\sum_{i=1}^{N-1}[z_N^2+z_1^2-2z_Nz_1+2\sigma_\varepsilon^2]$$

$$= \frac{(z_N-z_1)^2}{N-1} + \frac{2\sigma_\varepsilon^2}{N-1}$$

Now substitute the simplified quantities back into equation (B-4) to obtain

$$E\{\frac{\sum_{i=1}^{N-1}(x_i-\bar{x})^2}{N-2}\} = \frac{1}{N-2}\{\sum_{i=1}^{N-1}(z_{i+1}-z_i)^2+2(N-1)\sigma_\varepsilon^2 - \frac{2}{N-1}(z_N-z_1)^2 -$$

$$\frac{4\sigma_\varepsilon^2}{N-1} + \frac{(z_N-z_1)^2}{N-1} + \frac{2\sigma_\varepsilon^2}{N-1}\}$$

$$= \frac{1}{N-2}\{\frac{2N(N-2)}{N-1}\sigma_\varepsilon^2 + \sum_{i=1}^{N-1}(z_{i+1}-z_i)^2 -$$

$$\frac{(z_N-z_1)^2}{N-1}\}$$

$$E\{\frac{\Sigma(x_i-\bar{x})^2}{N-2}\} = \frac{2N}{(N-1)}\sigma_\varepsilon^2 + \frac{1}{(N-2)}[\sum_{i=1}^{N-1}(z_{i+1}-z_i)^2 - \frac{(z_N-z_1)^2}{N-1}]$$

(B-5)

In order to compare this estimate with that in equation (B-1) we must put it in the form of equation (B-3). Multiplying both sides of equation (B-5) by (N-1)/2N we obtain:

$$\frac{N-1}{2N}E\{\frac{\Sigma(x_i-\bar{x})^2}{N-2}\} = \sigma_\varepsilon^2 + \frac{N-1}{2N(N-2)}[\sum_{i=1}^{N-1}(z_{i+1}-z_i)^2-\frac{(z_N-z_1)^2}{N-1}]$$

Another estimator to be considered for estimating the variance is then

$$S_\varepsilon^2 = (N-1)\sum_{i=1}^{N}(x_i-\bar{x})^2/2N(N-2) \qquad \text{(B-6)}$$

Another procedure for estimating the variance is to use the sum of squared residuals obtained when fitting a linear model through the observations using a least squares regression technique. This estimate of the variance will contain a bias factor due to lack of fit of the model as discussed in Chapter I. Using the linear least squares regression equation for estimating the variance, the computation of the expected bias factor follows*.

---

*For derivation of the LLSR equations see Draper and Smith [5].

The LLSR equation for estimating the variance is:

$$MS_E = \{[\sum_{i=1}^{N} y_i^2 - (\sum_{i=1}^{N} y_i)/N^2] - \frac{\sum_{i=1}^{N}(t_i-\bar{t})(y_i-\bar{y})}{\sum_{i=1}^{N}(t_i-\bar{t})^2}[\sum_{i=1}^{N} t_i y_i - \frac{\sum_{i=1}^{N} t_i \Sigma y_i}{N}]\}/(N-2)$$

Compute the expected bias factor due to lack of fit of the LLSR model:

$$E(MS_E) = \frac{1}{N-2}\{[\sum_{i=1}^{N} y_i^2 - (\sum_{i=1}^{N} y_i)^2/N] - \frac{1}{\sum_{i=1}^{N}(t_i-\bar{t})^2}$$

$$[E(\sum_{i=1}^{N}(t_i-\bar{t})(y_i-\bar{y}) \sum_{i=1}^{N} t_i y_i)] \tag{B-7}$$

$$- \frac{1}{N\sum_{i=1}^{N}(t_i-\bar{t})^2}[E(\sum_{i=1}^{N}(t_i-\bar{t})(y_i-\bar{y})\sum_{i=1}^{N} t_i \sum_{i=1}^{N} y_i/N)]\}$$

Examine each term inside the brackets separately:

First term:

$$E[\sum_{i=1}^{N} y_i^2 - (\sum_{i=1}^{N} y_i)^2/N] = E[\sum_{i=1}^{N}(z_i^2 + 2z_i\varepsilon_i + \varepsilon_i^2) - \frac{1}{N}[(\sum_{i=1}^{N} z_i + \varepsilon_i)^2]$$

$$= \sum_{i=1}^{N} z_i^2 + N\sigma_\varepsilon^2 - \frac{1}{N}[N\sigma_\varepsilon^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} z_i z_j]$$

$$= \sum_{i=1}^{N} z_i^2 + (N-1)\,\sigma_\varepsilon^2 + (\sum_{i=1}^{N} z_i)^2/N$$

Second term:

$$\frac{1}{\sum_{i=1}^{N}(t_i-\bar{t})^2} = E[\sum_{i=1}^{N} (t_i-\bar{t})(y_i-\bar{y}) \sum_{i=1}^{N} t_i y_i] =$$

$$\sum_{i=1}^{N} \frac{1}{(t_i-\bar{t})^2} E[\sum_{i=1}^{N} (t_i z_i + t_i\varepsilon_i - \bar{t}z_i - \bar{t}\varepsilon_i - t_i\bar{z} - t_i\bar{\varepsilon} + \bar{t}\bar{z} + \bar{t}\bar{\varepsilon})$$

$$\sum_{i=1}^{N} (t_i z_i + t_i \varepsilon_i)$$

$$= \frac{1}{\sum_{i=1}^{N}(t_i-\bar{t})^2} [(\sum_{i=1}^{N} t_i z_i)^2 - (\bar{t} \sum_{i=1}^{N} z_i \sum_{i=1}^{N} t_i z_i) - \bar{z} \sum_{i=1}^{N} t_i \sum_{i=1}^{N} t_i z_i$$

$$+ N\bar{z}t \sum_{i=1}^{N} t_i z_i + \sum_{i=1}^{N} t_i \sigma_\varepsilon^2 - \bar{t} \sum_{i=1}^{N} t_i \sigma_\varepsilon^2 - \sum_{i=1}^{N} \sum_{j=1}^{N} t_i t_j (\frac{\sigma_\varepsilon^2}{N})$$

$$- \bar{t}\sigma_\varepsilon^2 \sum_{i=1}^{N} t_i]$$

$$= \frac{1}{\sum_{i=1}^{N}(t_i-\bar{t})^2} [(\sum_{i=1}^{N} t_i z_i)^2 - \bar{z} \sum_{i=1}^{N} t_i \sum_{i=1}^{N} t_i z_i + \sum_{i=1}^{N} t_i^2 \sigma_\varepsilon^2 -$$

$$\bar{t} \sum_{i=1}^{N} t_i \sigma_\varepsilon^2]$$

Third term:

$$\frac{1}{\sum_{i=1}^{N}(t_i-\bar{t})^2}\, E\left[\sum_{i=1}^{N}(t_i-\bar{t})(y_i-\bar{y})\sum_{i=1}^{N}t_i\sum_{i=1}^{N}y_i/N\right] =$$

$$\frac{1}{N(\sum_{i=1}^{N}t_i-\bar{t})^2}\, E\left[\sum_{i=1}^{N}t_iz_i + \sum_{i=1}^{N}t_i\varepsilon_i - \sum_{i=1}^{N}\bar{t}z_i - \sum_{i=1}^{N}\bar{t}\varepsilon_i -\right.$$

$$\left.\sum_{i=1}^{N}t_i\bar{z} - \sum_{i=1}^{N}t_i\bar{\varepsilon} + \sum_{i=1}^{N}\bar{t}\bar{z} + \sum_{i=1}^{N}\bar{t}\bar{\varepsilon}\right]\left[\sum_{i=1}^{N}t_i\sum_{i=1}^{N}z_i + \sum_{i=1}^{N}t_i\sum_{i=1}^{N}\varepsilon_i\right]$$

$$= \frac{1}{N\sum_{i=1}^{N}(t_i-\bar{t})^2}\left[\sum_{i=1}^{N}t_i\sum_{i=1}^{N}z_i\sum_{i=1}^{N}t_iz_i + (\sum_{i=1}^{N}t_i)^2\sigma_\varepsilon^{\,2} -\right.$$

$$\bar{t}(\sum_{i=1}^{N}z_i)^2\sum_{i=1}^{N}t_i - N\bar{t}\sum_{i=1}^{N}t_i\sigma_\varepsilon^{\,2} - \bar{z}(\sum_{i=1}^{N}t_i)^2\sum_{i=1}^{N}z_i - (\sum_{i=1}^{N}t_i)^2\sigma_\varepsilon^{\,2}$$

$$\left. + N\bar{t}\bar{z}\sum_{i=1}^{N}t_i\sum_{i=1}^{N}z_i + N\bar{t}\sum_{i=1}^{N}t_i\sigma_\varepsilon^{\,2}\right.$$

$$= \frac{1}{N\sum_{i=1}^{N}(t_i-\bar{t})^2}\left[\sum_{i=1}^{N}t_i\sum_{i=1}^{N}z_i\sum_{i=1}^{N}t_iz_i - \bar{t}\,(\sum_{i=1}^{N}z_i)^2\sum_{i=1}^{N}t_i\right]$$

Replacing the quantities in equation (B-7) with their expected values:

$$E(MS_E) = \frac{1}{N-2}\{\sum_{i=1}^{N} z_i^2 + (N-1)\sigma_\varepsilon^2 - (\sum_{i=1}^{N} z_i)^2/N - \frac{1}{\sum_{i=1}^{N}(t_i-\bar{t})^2}[(\sum_{i=1}^{N} t_i z_i)^2 \tag{B-8}$$

$$- \bar{z} \sum_{i=1}^{N} t_i \sum_{i=1}^{N} t_i z_i + \sigma_\varepsilon^2 \sum_{i=1}^{N} t_i^2 - \sigma_\varepsilon^2 \bar{t} \sum_{i=1}^{N} t_i]$$

$$+ \frac{1}{N \sum_{i=1}^{N}(t_i-\bar{t})^2}[\sum_{i=1}^{N} t_i \sum_{i=1}^{N} z_i \sum_{i=1}^{N} t_i z_i - \bar{t}\,(\sum_{i=1}^{N} z_i)^2 \sum_{i=1}^{N} t_i]\}$$

Since the trials are consecutive integer values, a closed form expression for t in terms of N can be made where

$$\sum_{i=1}^{N} t_i = \frac{N(N+1)}{2}$$

$$\sum_{i=1}^{N} t_i^2 = \frac{N(N+1)(2N+1}{6}$$

$$\bar{t} = \frac{N+1}{2}$$

$$\sum_{i=1}^{N}(t_i-\bar{t})^2 = \frac{N(N^2+1)}{12}$$

then rewriting equation (B-8) in terms of N where applicable

$$E(MS_E) = \frac{1}{N-2}\{\sum_{i=1}^{N} z_i^2 + (N-1)\sigma_\varepsilon^2 - \frac{(\sum_{i=1}^{N} z_i)^2}{N} - \frac{12}{N(N^2+1)}[(\sum_{i=1}^{N} t_i z_i)^2$$

$$- \frac{N(N+1)}{2}\bar{z}\sum_{i=1}^{N} t_i z_i + \sigma_\varepsilon^2 (\frac{N(N+1)(2N+1)}{6}) - \sigma_\varepsilon^2 (\frac{N+1}{N})(\frac{N(N+1)}{2})]$$

$$+ \frac{12}{N^2(N^2+1)}[\frac{N(N+1)}{2}\sum_{i=1}^{N} z_i \sum_{i=1}^{N} t_i z_i - (\frac{N+1}{2})(\frac{N(N+1)}{2})$$

$$(\sum_{i=1}^{N} z_i)^2]\}$$

Factoring and combining like terms:

$$E(MS_E) = \sigma_\varepsilon^2 [\frac{N^4-2N^3+N^2+N}{N^4-2N^3+N^2-2N}] + \frac{12}{N(N-2)(N^2+1)}[\frac{N^3+N}{12}\sum_{i=1}^{N} z_i^2 -$$

$$- \frac{(N^2+1)(\sum_{i=1}^{N} z_i)^2}{12} - (\sum_{i=1}^{N} t_i z_i - \frac{N+1}{2}\sum_{i=1}^{N} z_i)^2] \qquad \text{(B-9)}$$

Putting this in the form of equation (2-8) yields

$$\frac{N(N^2+1)(N-2)}{N(N^3-2N^2+N+1)} E[MS_E] = \sigma_\varepsilon^2 + \frac{12}{N(N^3-2N^2+N+1)}[\frac{N^3+N}{12}\sum_{i=1}^{N} z_i^2 -$$

$$- \frac{(N^2+1)(\sum_{i=1}^{N} z_i)^2}{12} - (\sum_{i=1}^{N} t_i z_i - \frac{N+1}{2}\sum_{i=1}^{N} z_i)^2]$$

Thus another estimator for the variance then would be

$$S_{\varepsilon}^{2} = (N^2+1)(N-2)MS_E/(N^3-2N^2+N+1) \qquad \text{(B-10)}$$

APPENDIX C

NONLINEAR TEST PROCEDURE

Estimation of Parameters and Variance

Step 1. Since we know that the degree of nonlinearity for the performance curve function is small, a linearization method for estimating the parameters should work quite well [10]. If a Taylor series expansion of the function $f(t_i,a,b)$ about the point (a,b) is carried out and curtailed after the first derivative, then an approximate estimate for the function $f(t_i,a,b)$ is then:

$$f(t_i,a,b) = f(t_i,a_o,b_o) + \left.\frac{\partial f(t,a,b)}{\partial a}\right|_{a_o}(a-a_o) + \left.\frac{\partial f(t,a,b)}{\partial b}\right|_{b_o}(b-b_o) \qquad \text{(C-1)}$$

The model could then be written as:

$$Y_i = f(t_i,a,b) + \frac{\partial f(t,a,b)}{\partial a} + \frac{\partial f(t,a,b)}{\partial b} + \varepsilon_i \qquad \text{(C-2)}$$

To solve for new estimates of the parameters, minimize the sum of squared errors. Let

$d_1 = (a-a_o)$

$d_2 = (b-b_o)$

$\bar{\theta} = (a,b)$

$$\varepsilon_i^2 = [y_i - f(t_i, a_o, b_o) - \frac{\partial f(t,a,b)}{\partial a}\Big|_{a_o} (a - a_o)$$

$$- \frac{\partial f(t,a,b)}{\partial b}\Big|_{b_o} (b - b_o)]^2$$

then minimizing ε_i^2 with respect to $\bar{\theta}$ we get

$$\frac{\partial \varepsilon_i^2}{\partial a} = 2\{\frac{\partial f(t_i, \bar{\theta})}{\partial a}| [y_i - f(t_i, \bar{\theta}_o) -$$

$$- \frac{\partial f(t_i, \bar{\theta})}{\partial a}\Big|_{\bar{\theta}_o} d_1 - \frac{f(t_i, \bar{\theta})}{\partial b}\Big|_{\bar{\theta}_o} d_2]\}$$

$$\frac{\partial \varepsilon_i^2}{\partial b} = 2\{\frac{\partial f(t, \bar{\theta})}{\partial b}\Big|_{\bar{\theta}_o} [y_i - f(t, \bar{\theta}_o) -$$

$$\frac{\partial f(t, \bar{\theta})}{\partial a}\Big|_{\bar{\theta}_o} d_1 - \frac{\partial f(t, \bar{\theta})}{\partial b}\Big|_{\bar{\theta}_o} d_2]\}$$

Then solve the two equations in two unknowns for the direction vector that will improve the initial estimates of the parameters

$$\frac{\partial f(t_i, \bar{\theta})}{\partial a \; \partial a}\Big|_{\bar{\theta}_o} d_1 + \sum_{i=1}^{n} \frac{\partial f(t_i, \bar{\theta})}{\partial a \; \partial b}\Big|_{\bar{\theta}_o} d_2 = \sum_{i=1}^{n} \frac{\partial f(t_i, \bar{\theta})}{\partial a}\Big|_{\bar{\theta}_o}$$

$$[y_i - f(t_i, \bar{\theta}_o)]$$

$$\sum_{i=1}^{n} \frac{\partial f(t_i,\bar{\theta})}{\partial a\, \partial b}\Big|_{\bar{\theta}_o} d_1 + \sum_{i=1}^{n} \frac{\partial f(t_i,\bar{\theta})}{\partial b\, \partial b}\Big|_{\bar{\theta}_o} d_2 = \sum_{i=1}^{n} \frac{\partial f(t_i,\bar{\theta})}{\partial b}$$

$$[y_i - f(t_i,\bar{\theta}_o)]$$

The minimum distance to move in the new direction D = (d_1,d_2) from $\bar{\theta}_o = (a_o,b_o)$ will be

$$\bar{\theta}_{new} = \bar{\theta}_o + v_{min}\bar{D} \tag{C-3}$$

where v_{min} is evaluated as follows:
Compute the sum of squared errors for

$$SS(\bar{\theta}_v) = \Sigma[y_i - f(t,\bar{\theta}_v)]^2$$

where $\bar{\theta}_v = \bar{\theta} + v\bar{D}$
and $v = 0, 1/2, 1$

Let $Q(v) = SS(\bar{\theta}_v)$

then

$$v_{min} = 1/2 + 1/4\ [Q(0)-Q(1)]/[Q(1)-2Q(1/2)+Q(0)]$$

Using the new estimate of the parameter vector, $\bar{\theta}_{new}$, in equation (C-3) as the next starting value, begin another iteration.

When the value of D = 0, or some small incremental value, δ = .000001, the best estimate is obtained for the parameters that minimizes the sum of squared errors. The corresponding variance would be

$$S_\varepsilon{}^2 = \frac{SS(\bar{\theta}_{i-1})}{N-2} \qquad (C-4)$$

where $\bar{\theta}_{i-1}$ represents the parameter values for the i^{th} iteration.

Step 2. Determine degree of nonlinearity. The procedure, as discussed earlier, entails using Beale's measure of nonlinearity to determine if the degree of nonlinearity of the performance function is less than .01/ $F_{\alpha,p,N-p}$. To compute the measure of nonlinearity, $\hat{N}_\theta$, the function is evaluated along with the corresponding tangent plane approximation, equation (3-34), at points $\bar{\theta}_w$ = (a,b) in the neighborhood of $\hat{\theta}$ = (a,b). Since there are two parameters, a reasonable design for considering points in the neighborhood of $\bar{\theta}$ would be a 3^2 design. In order to keep the distance between $\bar{\theta}_w$, w = 1, 2, 3, ..., 8 and $\hat{\theta}$ in proportion to the size of $\hat{\theta}$ as $\hat{\theta}$ changes, compute the upper and lower values of the parameters in the 3^2 design as $\hat{a} \pm$ %(a); $\hat{b} \pm$ %(b) for each parameter respectively. A reasonable percentage value would be between 3% and 10%.

Step 3. Confidence interval for parameters. If the degree of nonlinearity satisfies inequality (3-38) then an approximation for the linear theory confidence region

$$SS(\bar{\theta})-SS(\hat{\bar{\theta}}) \leq p[\frac{SS(\hat{\bar{\theta}})}{(N-p)}]F_{\alpha,p,N-p}[1 + \frac{N(p+2)}{(N-p)p} \hat{N}_{\theta}] \quad (C\text{-}5)$$

can be used to find the confidence interval for each parameter. A direct search procedure can be used to find the end points of the confidence interval for each parameter as follows:

(i) Holding parameter "a" at its least squares estimated value, search along the axis of "b" in both directions from b until equation (C-5) is satisfied. The smallest value for b and the largest value for b that satisfy equation (C-5) are the lower and upper confidence limits respectively for b.

(ii) Repeat procedure i, reversing the roles for the parameters.

Step 4. Confidence interval for the slope. If the joint confidence region for the parameters can be defined, then a confidence interval for the slope can be determined. Having determined the confidence interval for at least one parameter, say parameter "b". we can conduct a search at incremental points within this interval along the parameter

axis to obtain peripheral points of the joint parameter confidence region. The value of the slope for a given trial at any peripheral point corresponds to either an upper or lower limit for the slope at those particular values of (a,b). To determine the lower and upper limits of the slope at a given trial for the entire joint confidence region, select the smallest and largest value of the slope computed over the various periphery points. The smallest and the largest slope value correspond to the lower and upper confidence limits respectively for the slope at that given trial.

Appendix D consists of a computer program which estimates the parameters and σ_ε, tests the degree of nonlinearity and computes the confidence limits for the slope at time t_i.

APPENDIX D

COMPUTER PROGRAM

```
      PROGRAM NLNTEST(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
      DIMENSION TIME(50),VE(10,20),ALPHA(10),BETA(10),T(20),
     COBS(50)
      COMMON IX
C
C     BRACKET PARAMETER ESTIMATES AND READ IN LOWER VALUE FOR
C     EACH.  CHOOSE LARGE NUMBER FOR INITIAL SUM OF SQUARES
C     VALUE, SSC.
C
      READ*,N,A,B,P,F,(OBS(I), I=1,N)
      READ*,(TIME(I),I=1,N)
      READ*,(T(K),K=1,N)
C
C     THIS PROGRAM SOLVES FOR PARAMETERS "A", "B" BY
C     MINIMIZING THE SUM OF SQUARED ERRORS USING A
C     GRADIENT TYPE SEARCH PROCEDURE.
C
      PARMA=A
      PARMB=B
      DO 11 K=1,100
      F11=F12=F21=F22=0.0
      Q1=Q2=0.0
      DO 12 I=1,N
      F1=(0.0-(1.0/TIME(I)**PARMB))
      F2=(0.0-(PARMA/(TIME(I)**PARMB))*ALOG(1.0/TIME(I)))
      F11=F11+F1*F1
      F12=F12+F1*F2
      F21=F21+F2*F1
      F22=F22+F2*F2
      Q1=Q1+(OBS(I)-(1.0-PARMA/(TIME(I)**PARMB)))*F1
      Q2=Q2+(OBS(I)-(1.0-PARMA/(TIME(I)**PARMB)))*F2
   12 CONTINUE
C
C     SOLVE FOR ELEMENTS OFDIRECTION VECTOR THAT WILL
C     IMPROVE OUR ESTIMATES OF PARAMETERS "A" AND "B".
C     FIND "D1" AND "D2" BY SOLVING A 2X3 MATRIX.
C
      IF (F115 .GT. H) GO TO 13
      F115=F115*H
      Q15=Q15*H
   13 IF (F225 .GT. H) GO TO 14
```

```
      F225=F225*H
      Q25=Q25*H
   14 D1=Q15
      D2=Q25
C
C     FIND MAXIMUM DISTANCE, VMIN, TO PROCEED IN NEW
C     DIRECTION FROM CURRENT PARAMETER VECTOR TO GAIN
C     AN IMPROVEMENT IN MINIMIZING  SUM OF SQUARED ERRORS
C
      W=1.0
   65 A1=PARMA
      B1=PARMB
      A2=PARMA+(W*.5)*D1
      B2=PARMB+(W*.5)*D2
      A3=PARMA+W*D1
      B3=PARMB+W*D2
      QA1=QA2=QA3=0.0
      DO 15 I=1,N
      QA1=QA1+(OBS(I)-(1.0-A1/(TIME(I)**B1)))**2
      QA2=QA2+(OBS(I)-(1.0-A2/(TIME(I)**B2)))**2
      QA3=QA3+(OBS(I)-(1.0-A3/(TIME(I)**B3)))**2
   15 CONTINUE
      VAL1=QA1+QA3
      VAL2=2.0*QA2
      IF (VAL1 .EQ. VAL2) GO TO 18
      VMIN=0.5+.25*(QA1-QA3)/(QA3-2.0*QA2+QA1)
   18 AV=PARMA+VMIN*W*D1
      BV=PARMB+VMIN*W*D2
      BVV=0.0-5.0
      IF (BV .GT. BVV) GO TO 2
      GO TO 16
    2 QV=0.0
      DO 19 I=1,N
      QV=QV+(OBS(I)-(1.0-AV/(TIME(I)**BV)))**2
   19 CONTINUE
      VAL=QV-QA1
      IF (VAL .LT. .00001) GO TO 20
      W=W*.5
      WRITE(6,97)QV
   97 FORMAT(" "/"QV= ",F11.8)
      GO TO 65
   20 D11=D1
      D22=D2
      IF(D11 .GT. 0.0) GO TO 61
      D11=(0.0-1.0)*D11
   61 IF (D11 .GT. .000001) GO TO 62
      IF (D22 .GT. 0.0) GO TO 63
      D22=(0.0-1.0)*D22
   63 IF (D22 .LT. .000001) GO TO 16
```

```
      IF (PARMBB .EQ. PARMB) GO TO 16
   62 PARMA=PARMA+VMIN*W*D1
      PARMB=PARMB+VMIN*W*D2
      PARMBB=PARMB
      IF (PARMB .LT. 5.0) GO TO 11
      GO TO 16
   11 CONTINUE
   16 S=(QA1/(N-2))**0.5
      WRITE (6,17) PARMA,PARMB,S
   17 FORMAT(" "/"PARMA= ",F11.8,5X,"PARMB= ",
     CF11.8,5X,"STD DEV= ",F11.8)
C
C    CHECK DEGREE OF NON-LINEARITY USING BEALE"S METHOD
C
      SNUM=0.0
      SDEM=0.0
C
C     GENERATE COMBINATIONS OF THE PARAMETERS ABOUT THE
C     ESTIMATED VALUES
C
      B1=PARMB-0.08*PARMB
      DO 21 I=1,9
      ALPHA(I)=PARMA-0.08*PARMA
      BETA(I)=B1+(I-1)*0.08*PARMB
      IF (I .LE. 3) GO TO 21
      ALPHA(I)=PARMA+0.08*PARMA
      BETA(I)=B1+(I-4)*0.08*PARMB
      IF (I .LE. 6) GO TO 21
      ALPHA(I)=PARMA
      IF (I .EQ. 8) GO TO 24
      BETA(I)=B1+(I-7)*0.08*PARMB
      GO TO 29
   24 BETA(I)=B1+2.0*0.08*PARMB
   29 IF (I .LE. 8) GO TO 21
      BETA(I)=PARMB
   21 CONTINUE
C
C     TEST DEGREE OF NON-LINEARITY
C
      DO 22  I=1,9
      DO 23 J=1,N
      VE(I,J)=1.0-ALPHA(I)/(TIME(J)**BETA(I))
  23  CONTINUE
  22  CONTINUE
      DO 25 I=1,8
      DEM=0.0
      DO 26 J=1,N
      X=0.0-1.0/(TIME(J)**BETA(9))
      Y=0.0-ALPHA(9)*ALOG(1.0/TIME(J))*((1.0/TIME(J))**BETA
```

```
     C(9))
      Q=(VE(I,J)-(VE(9,J)+(ALPHA(I)-ALPHA(9))*X+(BETA(I)
     C-BETA(9))*Y))**2
      SNUM=SNUM+Q
      D=(VE(I,J)-VE(9,J))**2
      DEM=DEM+D
   26 CONTINUE
      SDEM=SDEM+DEM**2
   25 CONTINUE
      RESULT=P*(S**2)*(SNUM/SDEM)
      WRITE(6,27) SNUM,SDEM,RESULT
   27  FORMAT(" ","SNUM = ",F11.8,5X,"SDEM = ",F11.8,5X,
     C"RESULT= ",F10.8)
C
C  DETERMINE IF LINEAR THEORY RESULTS CAN BE USED AS APPROX
C  IAMATIONS FOR PARAMETERS "A" AND "B"
C
      CPVAL=0.01/F
      IF (RESULT .LE. CPVAL) GO TO 31
      WRITE(6,28)
   28 FORMAT(" ","DEGREE OF NON-LINEARITY TOO LARGE")
      GO TO 60
C
C     COMPUTE THE CONFIDENCE LIMITS FOR  THE SLOPE AT
C     SPECIFIED TIME
C
C
C     COMPUTE THE COMPARISON VALUE FOR THE CONFIDENCE TEST
C
   31 UCL=P*(S**2)*F*(1.0+((N*(P+2.0))/(P*(N-P))))*RESULT)
      SSE=(S**2)*(N-P)
C
C     DETERMINE LOWER AND UPPER BOUNDS ON PARAMETER "B".
C     THEN CONDUCT A SEARCH ALONG THIS INTERVAL TO DETER-
C     MINE UPPER AND LOWER BOUNDS ON THE SLOPE.
C
      RED=0.0005
      BL=PARMB
   34 BL=BL-RED
      PARJ=0.0
      DO 32 I=1,N
      CBL=1.0-PARMA/(TIME(I)**BL)
      SPBL=(OBS(I)-CBL)**2
      PARJ=PARJ+SPBL
   32 CONTINUE
      DIFFBL=PARJ-SSE
      IF (DIFFBL .GE. UCL)GO TO 33
      GO TO 34
   33 WRITE(6,35)BL
```

```
   35 FORMAT(" "/"BL= ",F11.8)
      BU=PARMB
   38  BU=BU+RED
      PARK=0.0
      DO 36 I=1,N
      CBU=1.0-PARMA/(TIME(I)**BU)
      SPBU=(OBS(I)-CBU)**2
      PARK=PARK+SPBU
   36 CONTINUE
      DIFFBU=PARK-SSE
      IF (DIFFBU .GE. UCL) GO TO 37
      GO TO 38
   37 WRITE(6,39) BU
   39 FORMAT(" "," BU= ",F11.8)
      M=(BU-BL)/.005+1
      DO 40 K=1,N
      SLVAL=99.0
      UVAL=0.0
      B3=BL-.005
      DO 54 J=1,M
      B3=B3+.005
C
C     COMPUTE LOWER LIMIT FOR "A" GIVEN PARTICULAR VALUE "B"
C
      AL=PARMA
      RED=0.005
   43 AL=AL-RED
      PARL=0.0
      DO 41 I=1,N
      CAL=1.0-AL/(TIME(I)**B3)
      SPAL=(OBS(I)-CAL)**2
      PARL=PARL+SPAL
   41 CONTINUE
      DIFFAL=PARL-SSE
      IF (DIFFAL .GE. UCL) GO TO 42
      GO TO 43
C
C     COMPUTE VALUE OF SLOPE FOR GIVEN "AL","B",T(K)
C
   42 SLL=(AL*B3)/(T(K)**(B3+1.0))
C
C     SAVE SMALLEST VALUE OF SLOPE OVER FULL INTERVAL OF "B"
C     LOWER INTERVAL OF "A" AND SPECIFIED TIME, "T(K)"
C
      IF (SLL .LT. SLVAL) GO TO 46
      GO TO 47
   46 SLVAL=SLL
      ALVAL=AL
      BLVAL=B3
```

```
   47 AU=PARMA
   50 AU=AU+RED
      PARU=0.0
      DO 48 I=1,N
      CAU=1.0-AU/(TIME(I)**B3)
      SFAU=(OBS(I)-CAU)**2
      PARU=PARU+SFAU
   48 CONTINUE
      DIFFAU=PARU-SSE
      IF (DIFFAU .GE. UCL) GO TO 49
      GO TO 50
C
C     COMPUTE VALUE OF SLOPE FOR GIVEN "AU","B","T(K)"
C
   49 SUL=(AU*B3)/(T(K)**(B3+1.0))
C
C     SAVE LARGEST VALUE OF SLOPE OVER INTERVAL "B",UPPER
C     INTERVAL OF "A", AND SPECIFIED TIME, "T(K)"
C
      IF (SUL .GT. UVAL) GO TO 53
      GO TO 54
   53 UVAL=SUL
      AUVAL=AU
      BUVAL=B3
   54 CONTINUE
C
C     PRINT OUT UPPER AND LOWER CONFIDENCE INTERVAL VALUES
C     FOR SLOPE, GIVEN TIME,"T(K)" ALONG WITH THE COR-
C     RESPONDING VALUES OF "A", AND "B"
      WRITE (6,55)K,T(K),SLVAL,ALVAL,BLVAL,UVAL,AUVAL,BUVAL
   55 FORMAT(" ","T(",I2,")= ",F4.1,4X,"SLVAL= ",F10.8,
     C5X,"ALVAL= ",F10.8,5X,"BLVAL= ",F10.8/16X,"UVAL= ",
     CF10.8,5X,"AUVAL= ",F10.8,5X,"BUVAL= ",F10.8)
   40 CONTINUE
   60 STOP
      END
```

REFERENCES

1. AR 1000-1, dated 5 November 1974, Basic Policies for Systems Acquisition by the Department of the Army Headquarters, Department of the Army, Washington, D. C.

2. AR 10-4, dated 30 December 1974, Organization and Functions of U.S. Army Operational Test and Evaluation Agency, Headquarters, Department of the Army, Washington, D. C.

3. Barnes, Ralph M. and Harold T. Amrine, "The Effect of Practice on Various Elements Used in Screwdriver Work," Journal of Applied Psychology, April 1942, pp. 197-209.

4. Beale, E. M. L., Confidence Regions in Non-Linear Estimation," Royal Statistical Society London Journal, Series B 22-23, 1960-1961, pp. 41-87.

5. Caplan, Stanley H. and Walton N. Hancock, "The Effect of Simultaneous Motions on Learning," Journal of Methods Time Measurement, September-October 1963, pp. 23-31.

6. Draper, N. R. and Smith, H., Applied Regression Analysis, John Wiley and Sons, Inc., 1966.

7. Guttman, Irwin and Meeter, Duane A., "On Beale's Measures of Non-linearity," Technometrics, Vol. 7, No. 4, dated November 1965.

8. Hancock, Walton M. and Prakash Sathi, Learning Curve Research on Manual Operations: Phase II, Industrial Studies, rev. ed., MTM Association for Standards and Research, Research Report 113A, Fair Lawn, N.J., 1969.

9. Hancock, W. M. and Sathe, P., "Learning Curve Research on Manual Operations," Methods Time Measurement Research Studies, Report 113A, dated 1969.

10. Hartley, H. O., "The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares," Technometrics, Vol. 3, No. 2, May 1961, pp. 269-280.

11. Hoel, Paul G., Sidney C. Port and Charles J. Stone, Introduction to Statistical Theory, The Houghton Mifflin Company, Boston, 1971.

12. Meister, David, <u>Behavioral Foundations of System Development</u>, U.S. Navy Personnel Research and Development Center, Copywrite 1976, John Wiley and Sons, Inc.

13. Meyer, Paul L., <u>Introductory Probability and Statistical Applications</u>, The Addison-Wesley Publishing Company, 1970.

14. Schwartz, Shula, <u>The Learning Curve and Its Applications</u>, Chance Vought Aircraft, Inc., Library Bibliography 20, Dallas, Texas.